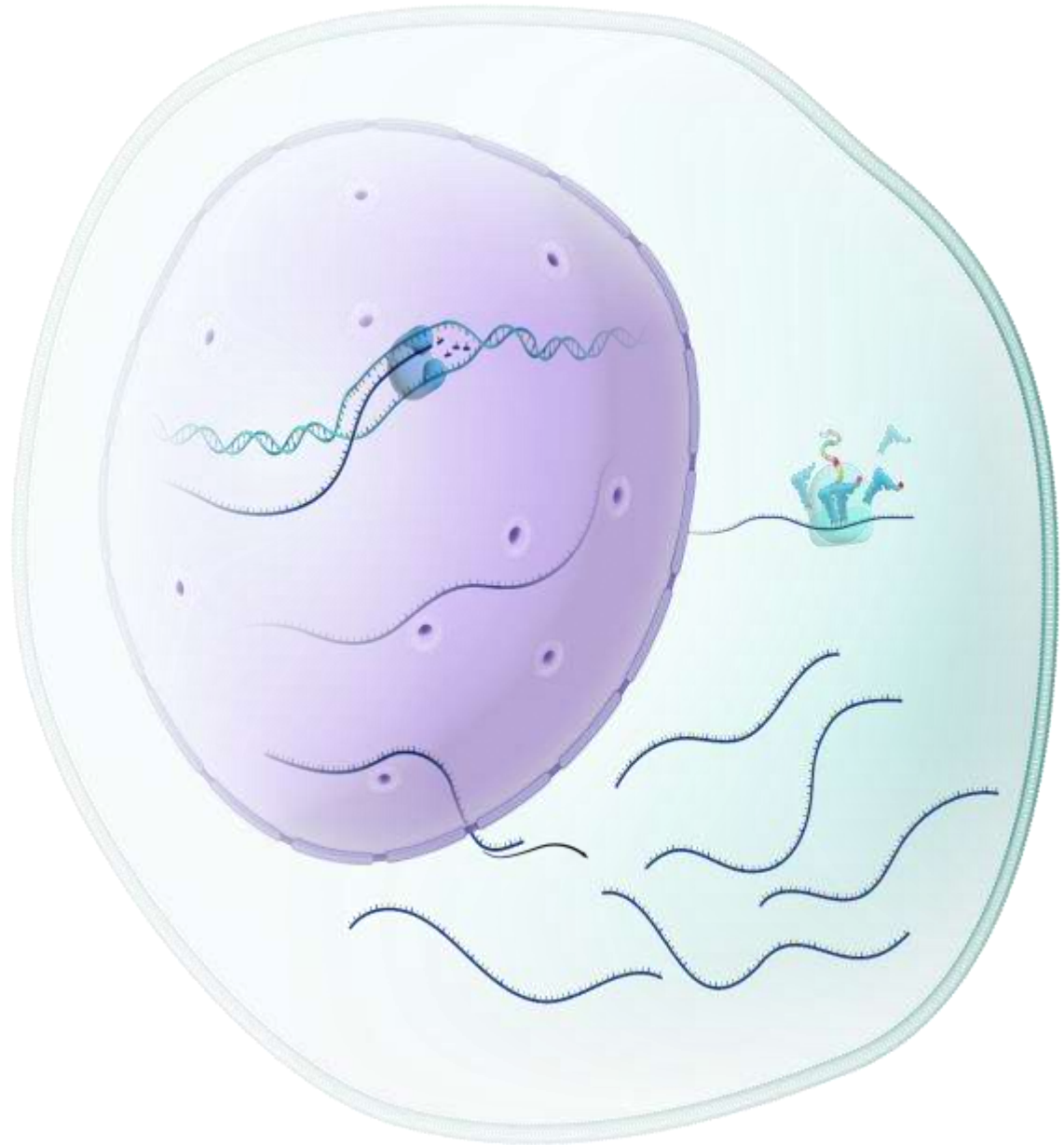


Regulation of gene transcription



Content

- 1. Introduction**
- 2. Prokaryotic regulation of transcription**
- 3. Operon model**
 - Lac operator
 - Lac structural genes
 - Lac repressor
- 4. Negative regulation of lac Operon**
 - Constitutive and Uninducible lac mutation
- 5. Positive regulation and catabolite repression**
- 6. Tryptophan Operon system**
 - Transcriptional attenuation
 - RNA binding protein mediated attenuation
- 7. Riboswitches**
- 8. Bacteriophage lambda- A transcriptional switch**
- 9. Eukaryotic regulation of transcription**
 - Influence of chromatin structure on transcription
 - DNA methylation and gene regulation
 - Post transcriptional gene regulation
 - RNA interference
 - miRNAs mediated RNAi
 - siRNAs mediated RNAi
 - piRNA

Prokaryotic transcriptional regulation is accomplished by gene regulatory proteins which binds with regulatory sequences located near the beginning of transcription units

Gene regulatory proteins= the product of regulatory genes

Acts as

Activator

Repressor

Activator protein + target DNA site (located near the promoter)= increase in rate of transcription

Such instances= **Positive Regulation**

Repressor protein + Target DNA site (the operator, located within the promoter)= prevention of gene expression (decreased rate of transcription/ no transcription)

Such instances= **Negative Regulation**

Thus gene regulation may be Positive or Negative

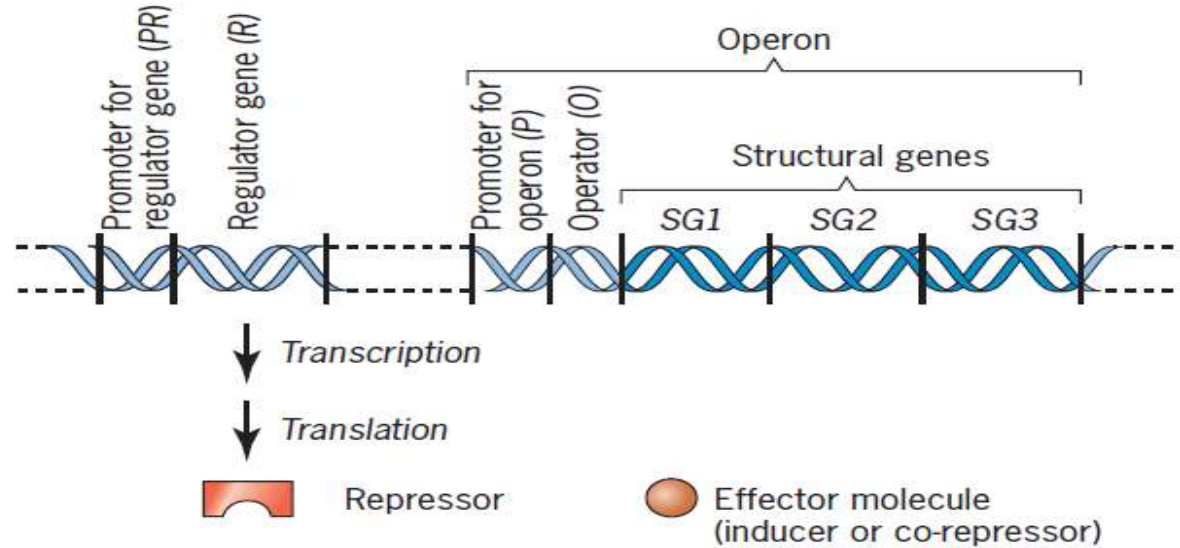
Negative regulation- mediated by **Repressor** that block or turn off transcription

Positive regulation- mediated by **Activator** required for the increased rate of transcription

Operon model:

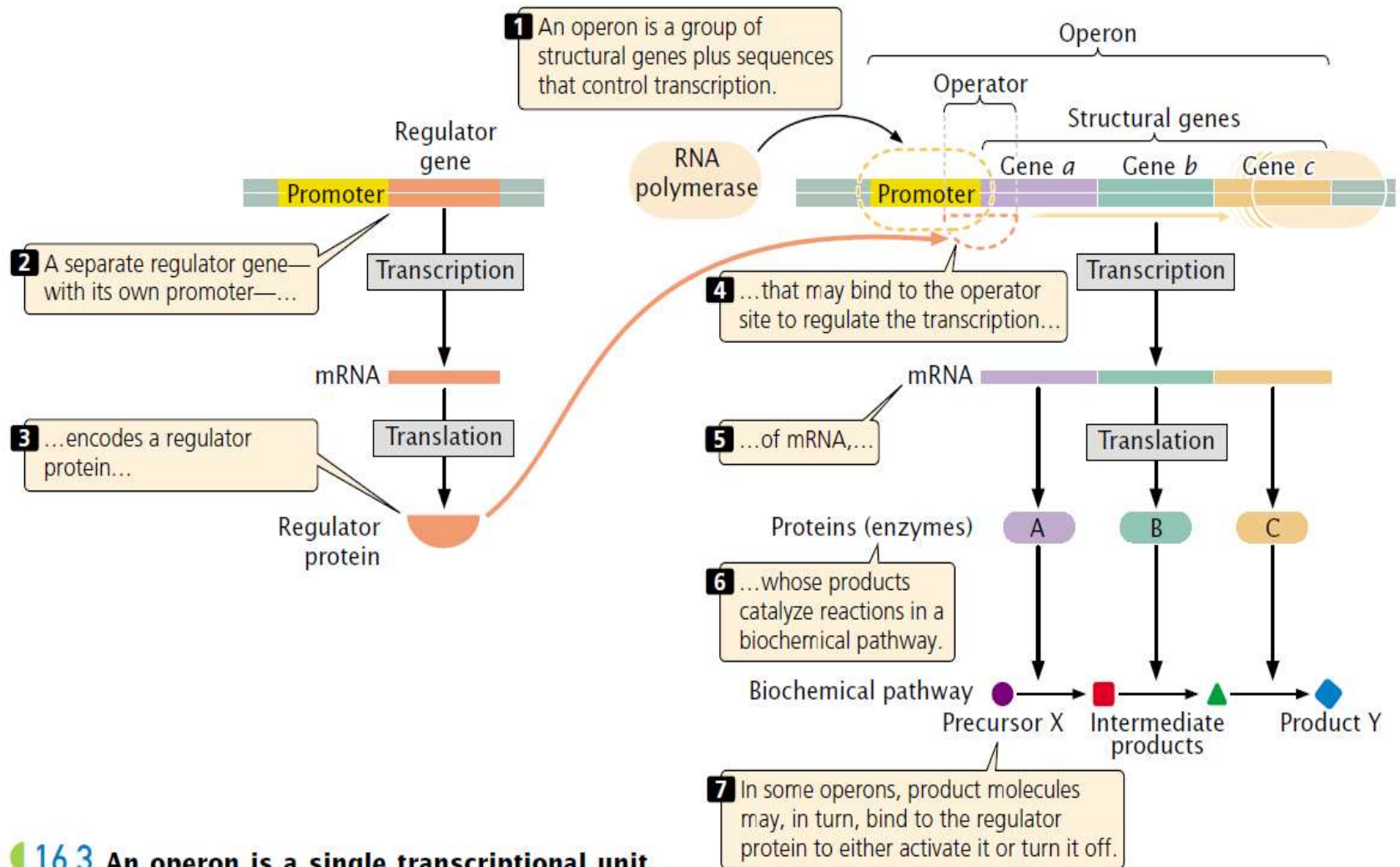
- Gene regulation occurs at the level of transcription in Bacteria
- Basic concept of gene regulation was explained by classical model called **Operon model**
- It was formulated by **Jacob** and **Monod** in 1961

The operon: components

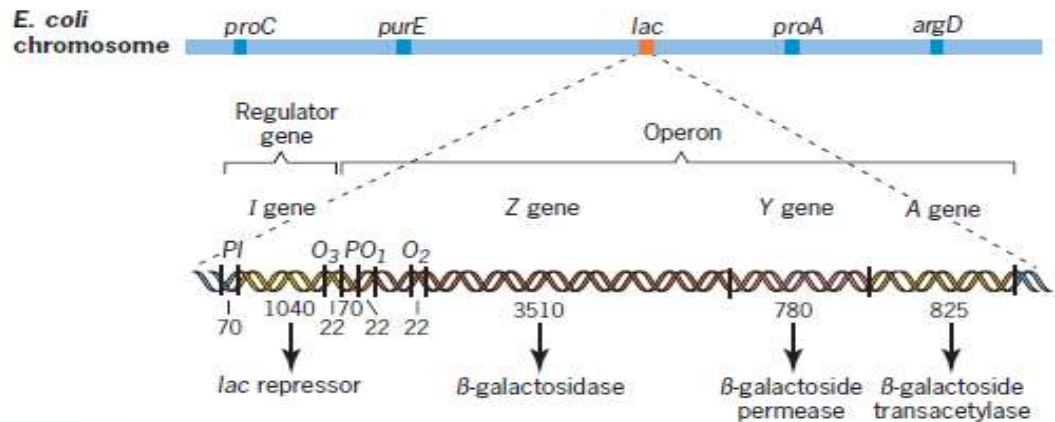


Operon = Unit of bacterial gene expression and regulation

- Includes **Structural Genes** and **Regulatory Sequences** recognized by regulatory gene product(s)
- All the genes in an Operon are expressed as a single unit
- This type of arrangement is common in prokaryotic genomes
- In this arrangement, single **Promoter** is required to initiate and regulate transcription of all structural genes
- An Operon consists of cis-acting regulatory elements and protein (or RNA) encoding structural genes
- A typical example of an Operon is **lactose (*lac*) Operon**
- Lac Operon contains,
 - **3 structural genes**
 - An **Operator**
 - A **Promoter**
 - The product of lac Operon is involved in conversion of the disaccharide sugar **Lactose** into its monosaccharide units **Glucose** and **Galactose**



16.3 An operon is a single transcriptional unit that includes a series of structural genes, a promoter, and an operator.



■ **FIGURE 18.5** The *lac* operon of *E. coli*. The *lac* operon consists of three structural genes, *Z*, *Y*, and *A*, plus the promoter (*P*) and three operators (*O₁*, *O₂*, and *O₃*). The regulator gene (*I*) is contiguous with the operon in the case of *lac* and has its own promoter (*P_I*). The numbers below the various genetic elements indicate their sizes in nucleotide pairs.

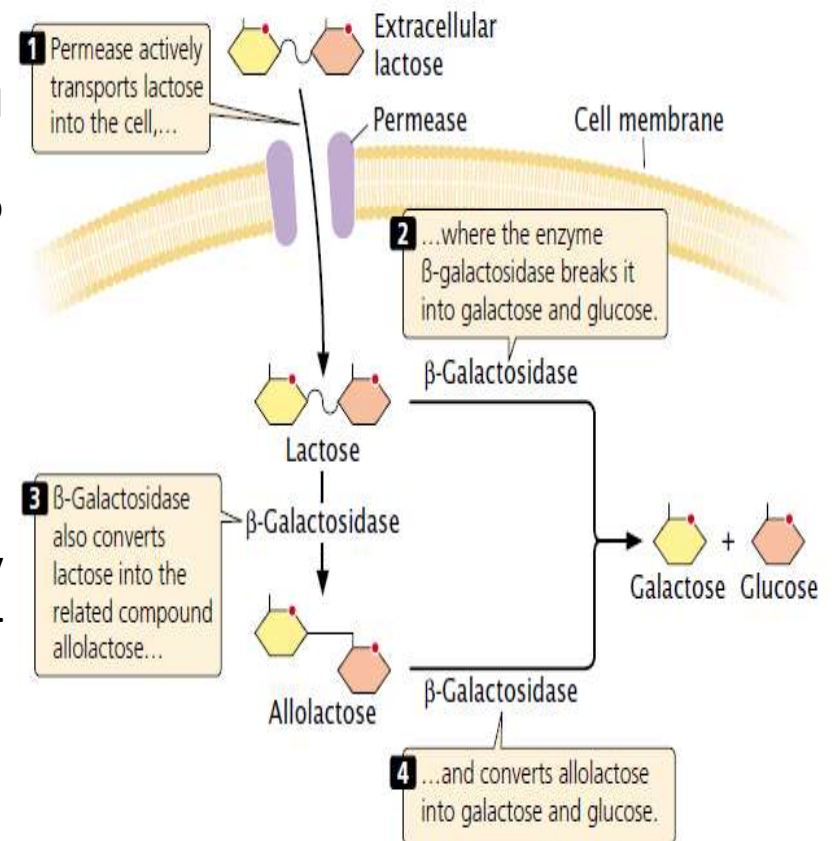
Lac operator:

- Short region of DNA
- Lies between the promoter and structural genes
- Interacts with a regulatory protein
- Controls the transcription of the Operon
- Lac Operon has 2 types of operators:
 1. **Primary operator- O_1**
 2. **Secondary or pseudo operator (O_2 and O_3)**
- **Primary lac operator:** palindromic sequence of ~26bp from -5 to +21
 - Cis acting site on the DNA to which the lac repressor binds
 - **Lac Repressor+ Operator (Primary Operator) = Prevents RNA Polymerase From Initiating transcription at the promoter**
- **Secondary operator:** located on either side of primary operator
 - O_2 centered at +410 position, downstream of transcription start site
 - O_3 centered at -90 position in upstream region
 - These are weak and have lesser affinity for Lac repressor

Lac structural genes:

- **Structural genes:** any genes that codes for a protein (or RNA)
- Lactose is a β - galactoside that *E coli* can use for energy and as a carbon source
- Lactose uptake and degradation is mediated by the products of three structural genes
- The roles of three structural genes are:
 - **Lac Z:**
 - Codes for enzyme β - galactosidase
 - Cleaves the lactose molecule to yield **Glucose** and **Galactose**
 - This enzyme also converts **Lactose** to **Allolactose**
 - **LacY:**
 - Codes for β - galactoside permease
 - Transports lactose into the cell
 - **LacA :**
 - Codes for β - galactoside transacetylase
 - Play role in detoxification of compounds by transferring an acetyl group from **acetyl-CoA** to β - galactosides

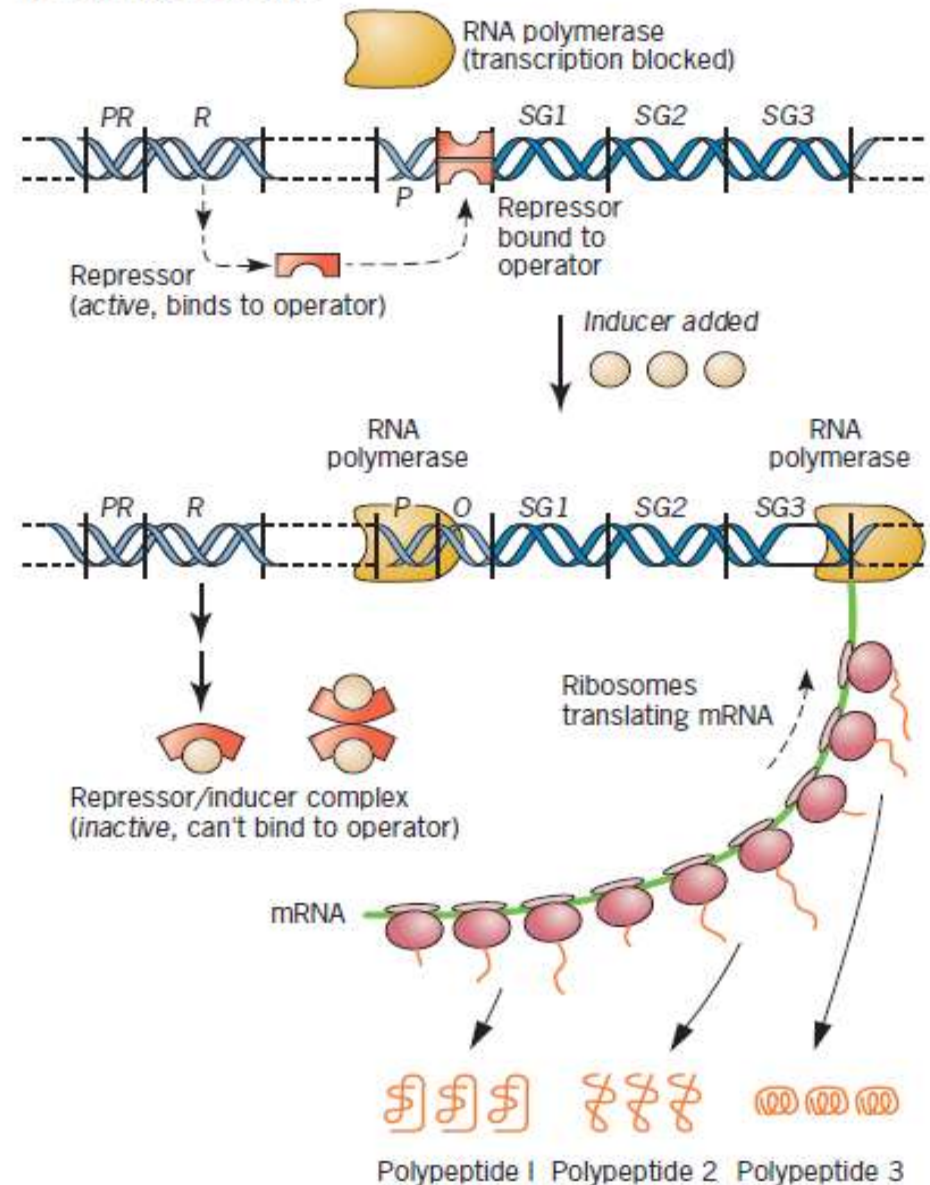
16.6 Lactose, a major carbohydrate found in milk, consists of 2 six-carbon sugars linked together.



Inducible Operon:

- The Operon which is expressed in the presence of a substance termed **Inducer**
- **Inducer**= Small molecule that triggers gene transcription by binding to a regulatory protein
- Lac Operon in *E coli* is an example for inducible Operon
- In lac Operon, **Lactose** and its analogs are inducer
- This lead to the expression of ***lac*** genes
- The relief of repression for Operon is termed as **Induction**

The operon: induction

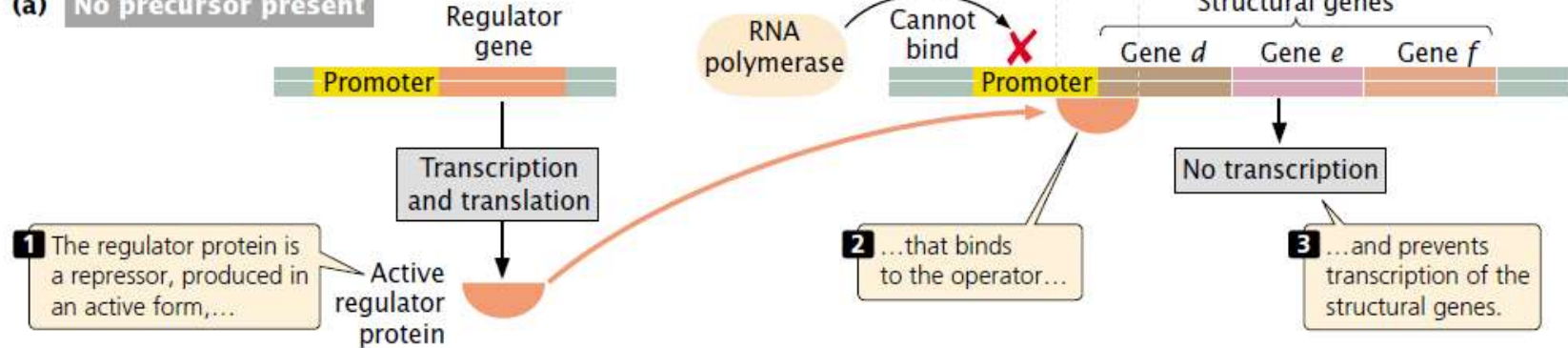


Lac repressor:

- Is the product of **trans-acting regulatory gene**
- Regulates transcription of the structural genes comprising lac Operon
- A **Regulatory Gene (I Gene)**- gene that codes for a protein (or RNA) involved in regulating the expression of other genes
- This gene encodes the **lac repressor proteins**
- Lac repressor molecule is a homotetramer made up of 4 identical subunits with molecular weights of ~38000
- It has 2 recognition sites:
 - **DNA binding site**- which recognize the specific **Operator** DNA sequence for the lac Operon; It only binds with only one DNA sequence with high affinity in the entire ***E coli*** genome- the **Lac Operator**
 - By binding to operator, the repressor prevents transcription initiation by **RNA polymerase** that has bound to its lac promoter site
 - **Allosteric site**- binds the lactose allosteric effector and similar molecules (analogs of lactose)
 - When the repressor binds to lactose or its analogs, the protein undergoes a conformational changes
 - This slight alteration in shape changes the DNA-binding site so that the repressor no longer has high affinity for the operator thus in response to binding lactose, the repressor falls off the DNA

Negative regulation of lac Operon:

(a) No precursor present



In the absence of **Inducer** (lactose/ analogs of lactose)

Lac Repressor binds to the **Lac Operator Site** and prevents transcription of the lac Operon

Binding of repressor prevents transcription
It acts as **Negative Regulator** and acts as a roadblock on the DNA

Effect of lac repressor binding to the operator:

Blocks the progression of RNA polymerase transcription

No β -galactosidase, no β -galactoside permease no transacetylase in the cell

All structural genes of lac Operon (**Z**, **Y** and **A gene**) are repressed

Gratuitous inducer:

The presence of **Inducer** (lactose or analogs)



Binds to allosteric site of the Lac repressor



Inactivates the operator binding site of the lac repressor protein



However lactose itself does not have this effect
Allolactose does

Allolactose is the actual inducer of the lac Operon
Formation of allolactose is catalyzed by β -galactosidase



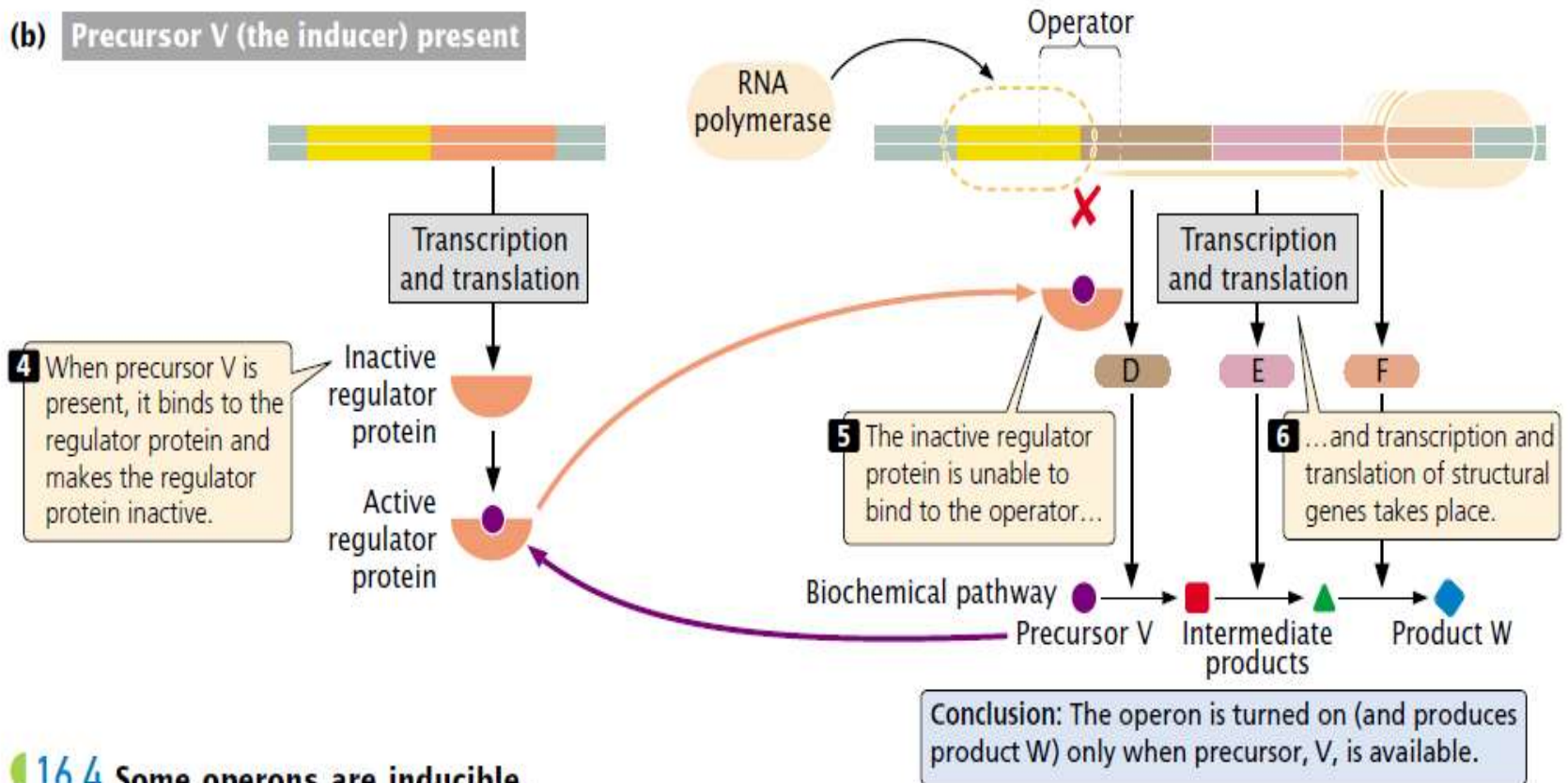
The lactose analog that is capable of inducing the lac Operon while not itself serving as a substrate for β galactosidase is known to be **Gratuitous Inducer**

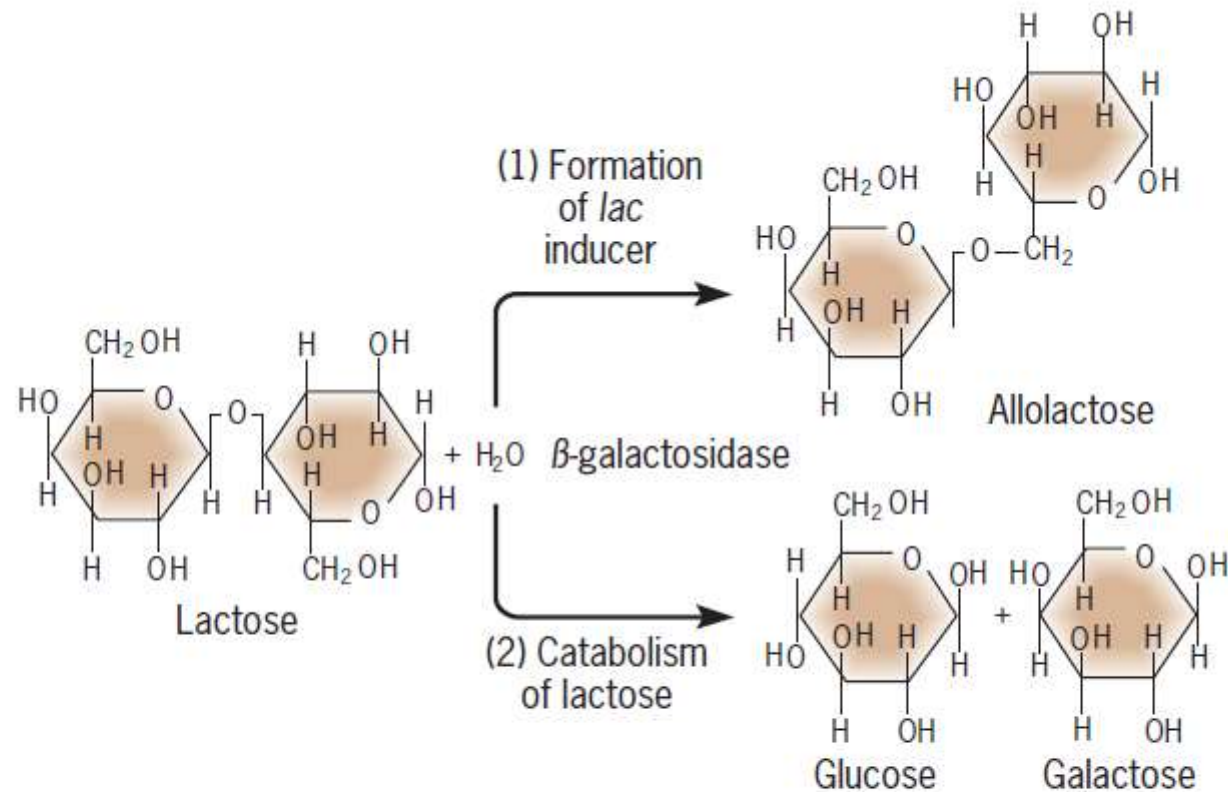


Example of gratuitous inducer: **Isopropylthiogalactoside (IPTG)**, A synthetic non- metabolizable analog of the lactose



Inactivation of lac repressor permits the induction of transcription of the structural genes of the lac Operon and through the translation of polycistronic mRNA
The enzymes β galactosidase, β galactoside permease and transacetylase appear in the cell in the coordinated fashion

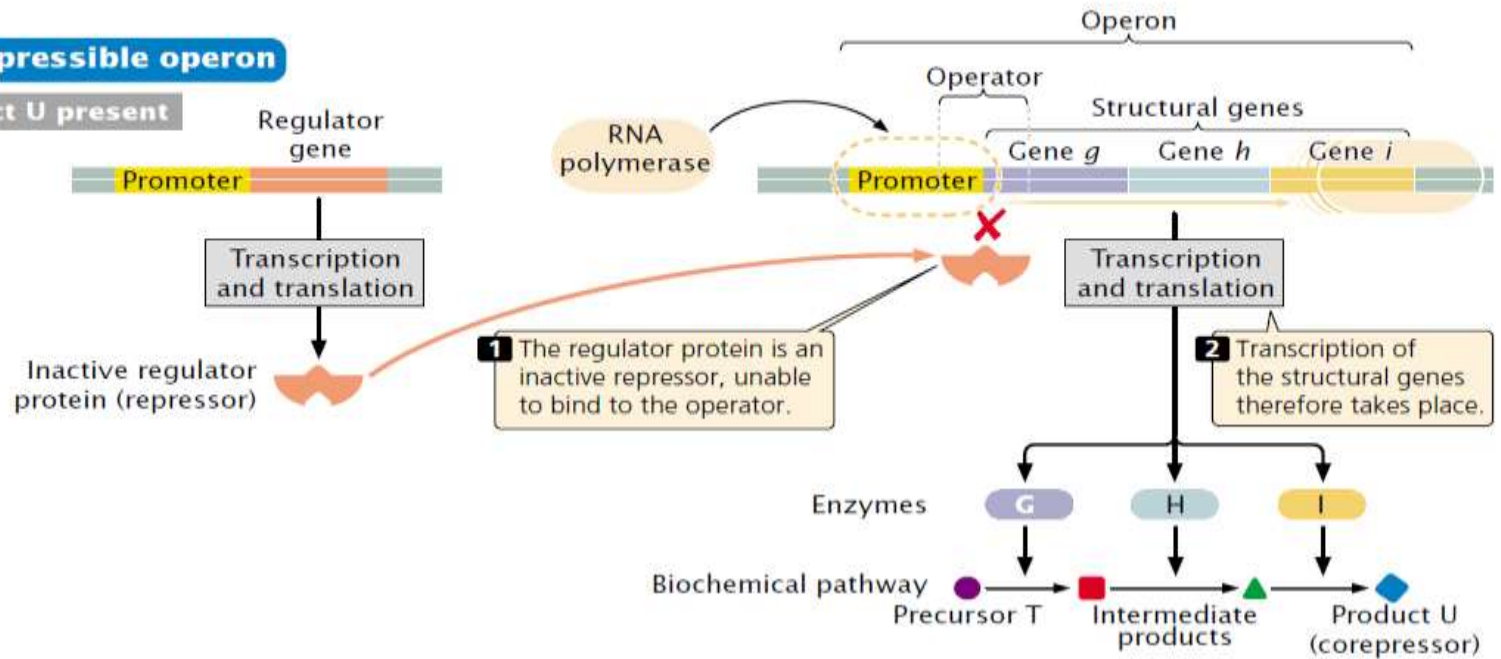
(b) Precursor V (the inducer) present**16.4 Some operons are inducible.**



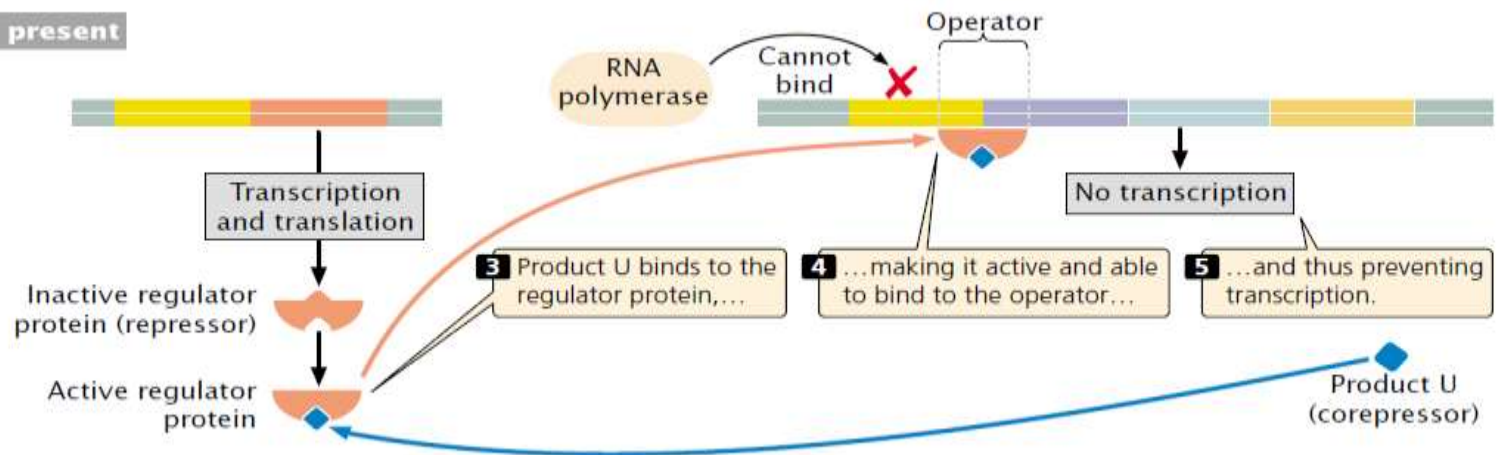
■ **FIGURE 18.6** Two physiologically important reactions catalyzed by β -galactosidase: (1) conversion of lactose to the *lac* operon inducer allolactose, and (2) cleavage of lactose to produce the monosaccharides glucose and galactose.

Negative repressible operon

(a) No product U present

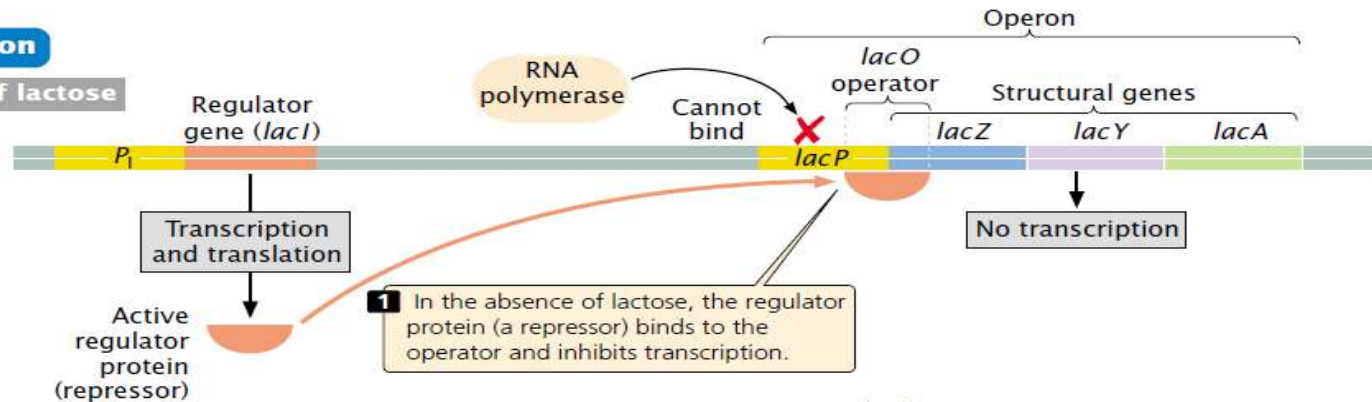


(b) Product U present

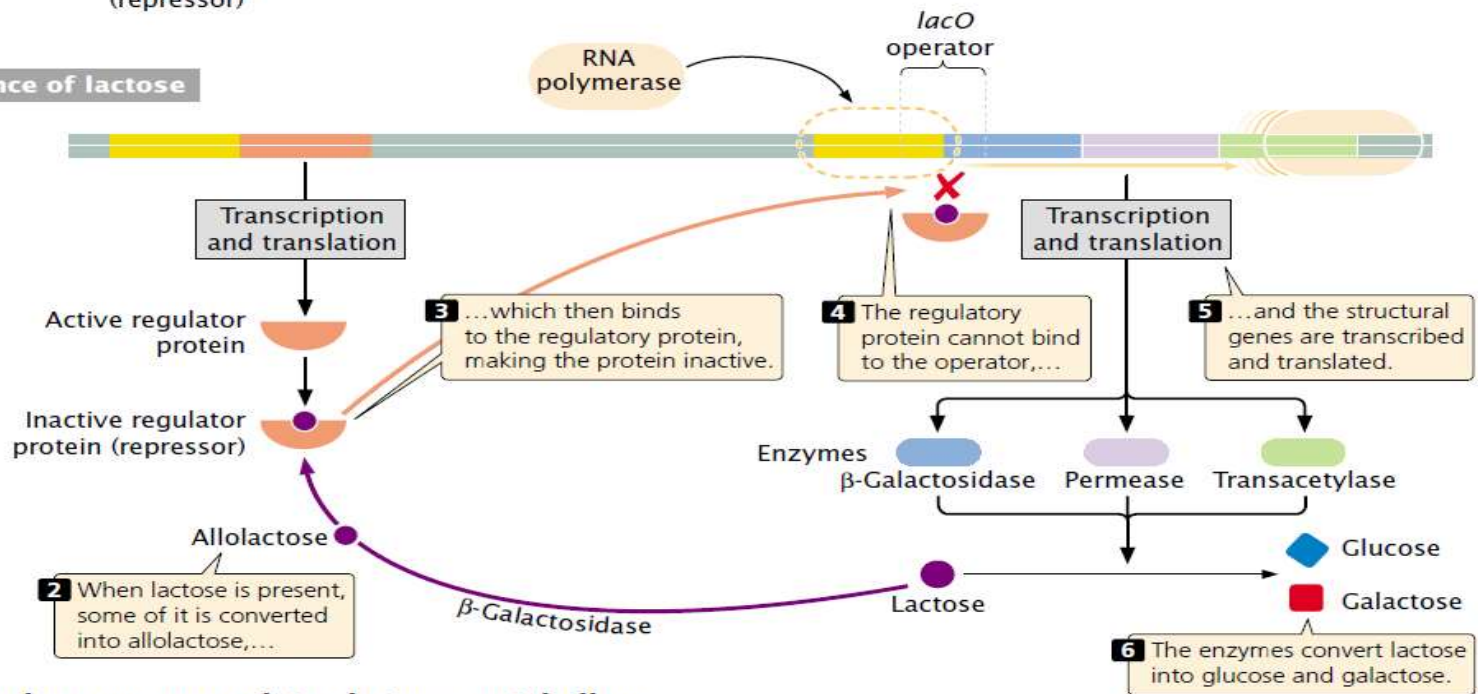


The *lac* operon

(a) Absence of lactose



(b) Presence of lactose



16.7 The *lac* operon regulates lactose metabolism.

Constitutive and uninducible lac mutation:

- A mutant, in which genes of an Operon are always transcribed even in the absence of inducer is called **constitutive mutant**

In constitutive lac mutation,

- All the three lac Operon genes are transcribed at all times
- That is, they are not turned off even in the absence of lactose (the expression of lac Operon genes cannot be repressed)
- This type of mutation is common with **negatively regulated operons**

Uninducible mutant: In this type of mutant, the expression of lac Operon genes cannot be induced

Lac mutations:

- Can be dominant or recessive
- Whether a particular lac mutation is dominant or recessive was determined by introducing wild type lac Operon into a strain with lac mutation
- Ex. ***lacI*** gene coding for **Repressor**
- ***lacI*⁺** = Wild type; ***lacI*⁻** = mutant type
- If mutant ***lacI*⁻** is recessive,
 - It means that the introduction of a normal ***lacI*⁺** gene in the cell restores control, irrespective of the presence of the mutant ***lacI*⁻** gene
- Most lac mutation are recessive to wild type
- Lac mutations may be **cis** or **trans** acting also
 - Recessive mutations that cannot be complemented are cis acting
 - Cis acting mutations affects only the expression of genes on the same DNA where it occurs

***lacI*⁻ mutation:**

- This type of mutants do not synthesize repressor or synthesize repressor that are not capable of binding to the operator
- *E coli* with ***LacI*⁻** mutant exhibits constitutive expression of the lac Operon
- ***lacI*⁻** mutant is generally recessive to wild type and trans acting

***Lac I*^d mutation:**

- This mutation is possible because the ***lacI*** polypeptide forms a homotetramer
- Monomers associate at random in the cell to form the active protein
- Mixture of normal and defective subunits can be non-functional causing the constitutive like the ***lacI*⁻** phenotype
- It is called by the generic names **dominant negative**
- ***Lac I*^d** mutations occur in the DNA binding site of the polypeptide
- ***LacI*^d** mutant is dominant, trans acting and constitutive

***lacI*^s mutation:**

- This type of mutants have changes in the repressor so that it cannot bind the inducer
- This type is dominant, trans acting and uninducible mutation

***Lac O*^c mutation:**

- It is a dominant cis acting and constitutive type of mutation
- A mutation in lac operator does not allow the binding of a normal repressor molecule with operator

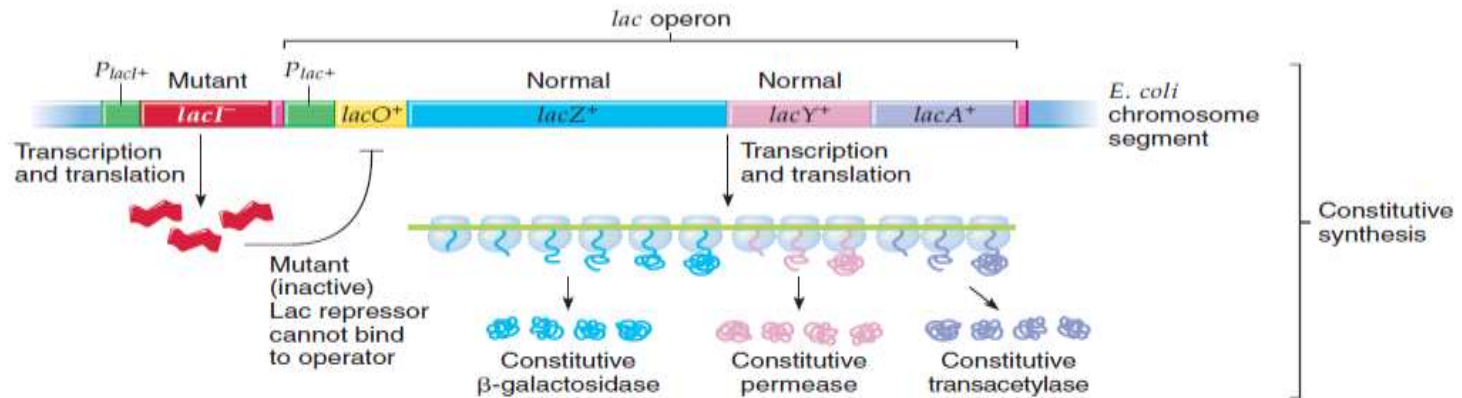
Lac P mutation:

It is recessive cis acting and uninducible mutation

Figure 17.9

Effects of a *lacI*⁻ mutation in a *lacI*⁻ *lacO*⁺ *lacZ*⁺ *lacY*⁺ haploid cell, and in a *lacI*⁺ *lacO*⁺ *lacZ*⁻ *lacY*⁺/*lacI*⁻ *lacO*⁺ *lacZ*⁺ *lacY*⁻ partial diploid strain of *E. coli*. (The *lacZ*⁻ and *lacY*⁻ mutations are missense mutations.)

- a) Haploid strain (in presence or absence of inducer). The mutant Lac repressor cannot bind to the *lacO*⁺ operator, resulting in constitutive synthesis of *lac* operon enzymes.



- b) Partial diploid in the absence of inducer. The *lacI*⁺ operon produces wild-type Lac repressors, whereas the *lacI*⁻ operon produces inactive Lac repressors. The mutant Lac repressor cannot bind to the *lacO*⁺ operators, but the wild-type Lac repressors can, so no transcription occurs from either operon.

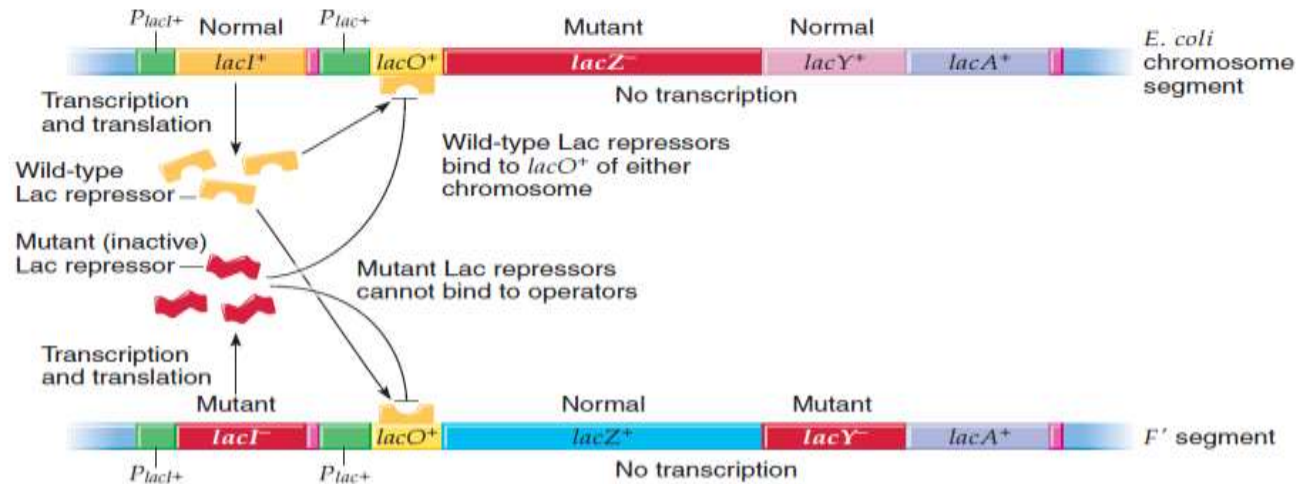
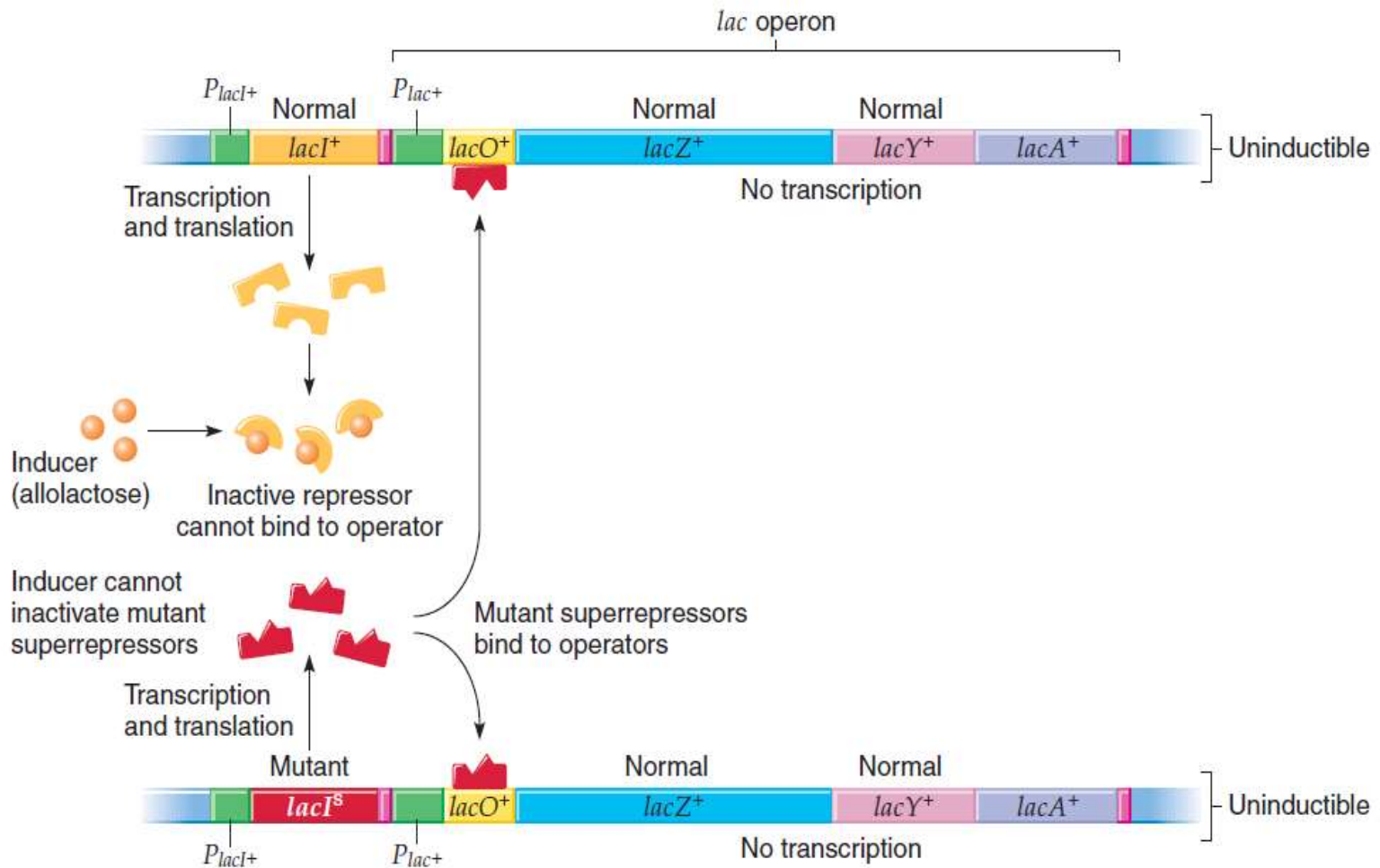


Figure 17.10

Dominant effect of *lacI^S* mutation over wild-type *lacI⁺* in a *lacI⁺ lacO⁺ lacZ⁺ lacY⁺ lacA⁺ / lacI^S lacO⁺ lacZ⁺ lacY⁺ lacA⁺* partial diploid cell growing in the presence of lactose.



Positive regulation and catabolite repression:

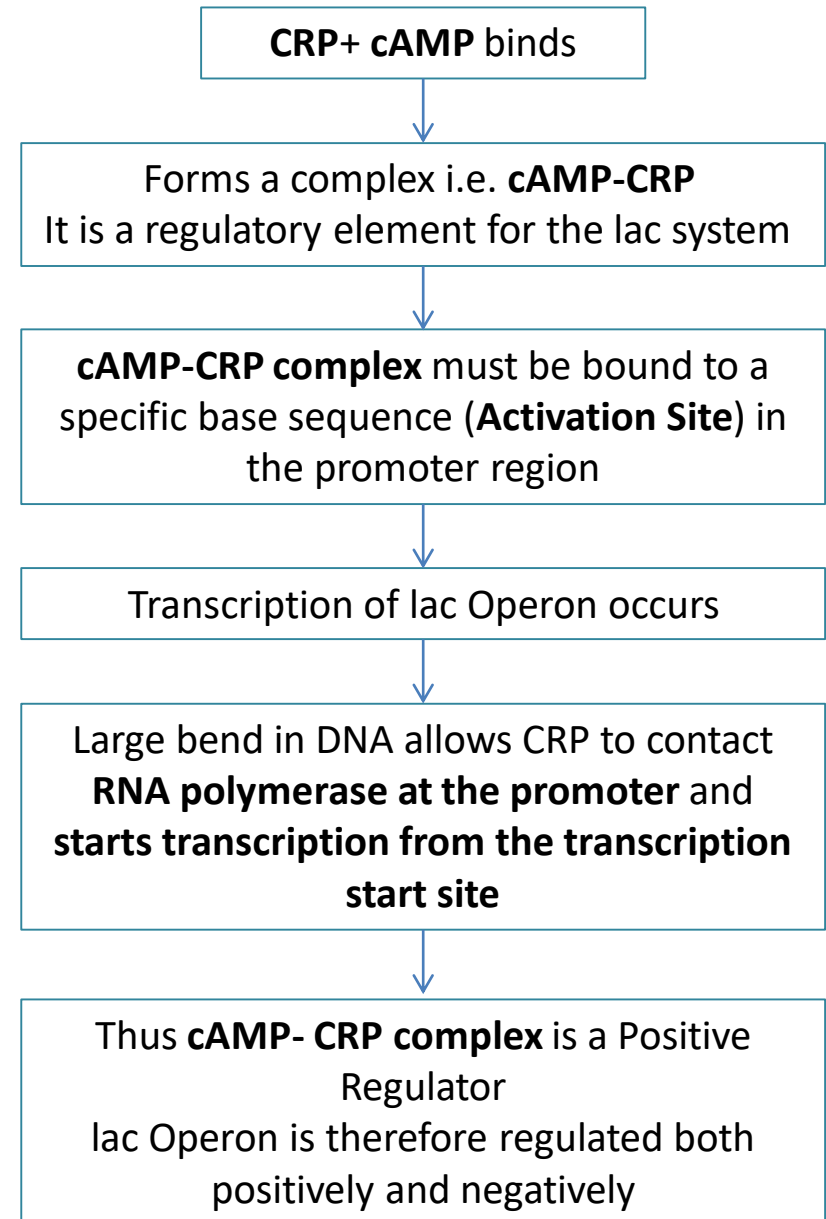
- When **Glucose** is available as an energy source, it is used in preference to other sugars
- For example,
 - In *E coli*, when it finds both **Glucose** and **Lactose** in the medium, it metabolizes the **Glucose** and represses the use of **Lactose**
 - The ability of glucose to inhibit the synthesis of certain enzymes, referred to as the **Glucose Effect**
- This was recognized by **Monod** in bacteria
- He observed,
 - When *E coli* encounters both Glucose and Lactose,
 - It metabolizes the Glucose first and represses the use of Lactose resulting in **Diauxic Growth**
 - This phenomenon is due to the **Repressive Effect of glucose** on the synthesis of enzymes required for the metabolism of other sugars
 - The term **Catabolite Repression** was introduced as a general name for Glucose Effect

Positive control system of transcription

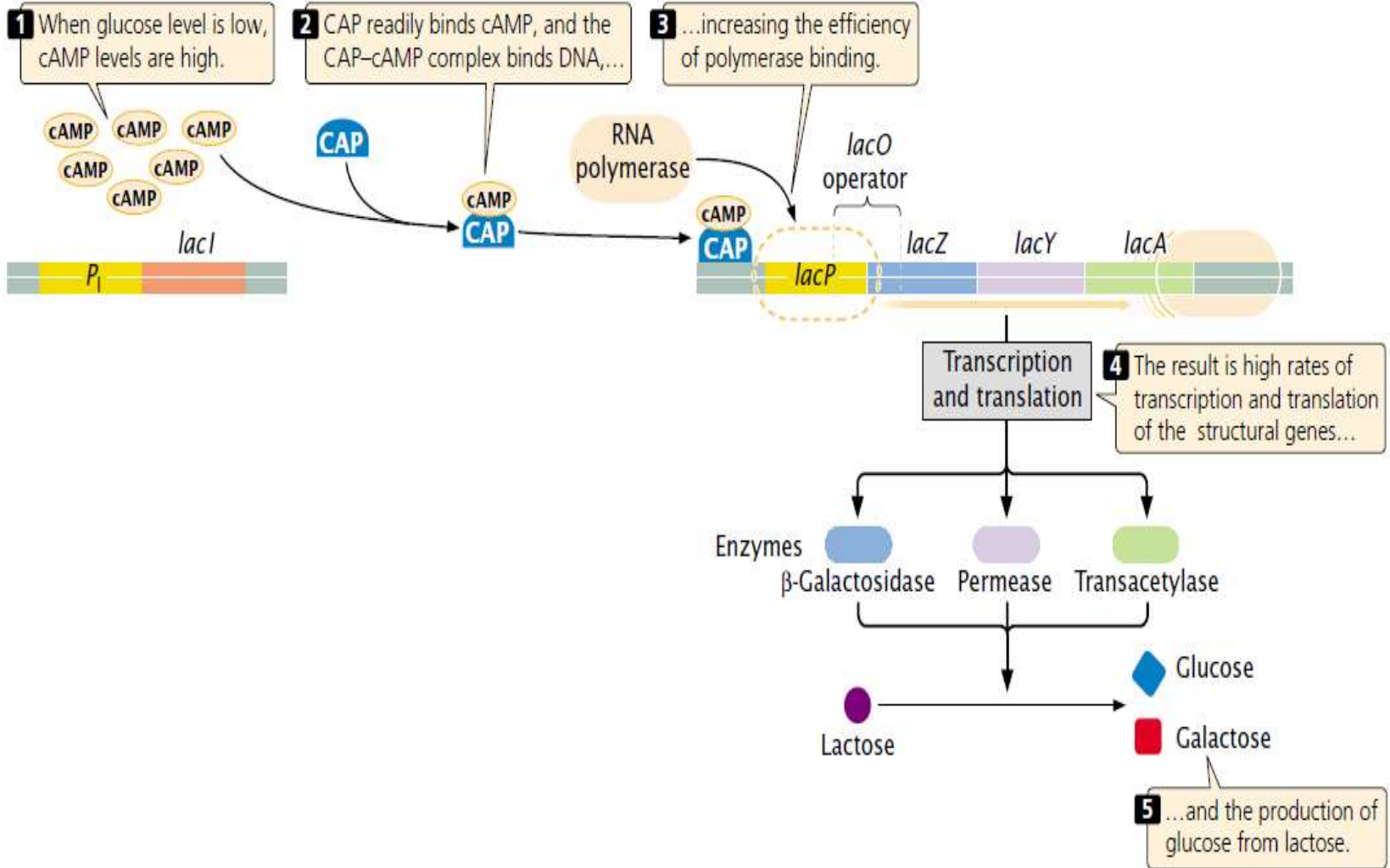
- By **cAMP** and the **cyclic AMP receptor protein**
- **CRP** –(also called the **Catabolite Gene Activator Protein, CAP**)
- Catabolite repression is caused by a **Reduction of the level of Intracellular cAMP**
- **CRP** – Homodimer and acts as **Positive Regulator Protein**
- It is activated by binding of single molecule of **cyclic AMP**
- It is needed for RNA polymerase to initiate transcription of many Operon of *E coli*
- **cAMP** is synthesized by the enzyme **Adenylate cyclase** and its concentration is related to glucose concentration
- The **Glucose Level** in the cell is **high= cAMP concentration is low**
- The **Glucose Level** is low= **cAMP concentration is high**

E coli contains a protein called **cAMP receptor protein (CRP)** that is encoded by a gene called *crp*

Mutants of either *crp* or the **Adenylate Cyclase** gene are unable to synthesize **lac mRNA** which indicates that both CRP function and cAMP are required for lac mRNA synthesis



When glucose is low



When glucose is high

1 When glucose level is high, cAMP levels are low, and cAMP is less likely to bind to CAP.

cAMP cAMP

CAP

CAP

2 RNA polymerase cannot bind to DNA as efficiently;...

RNA polymerase

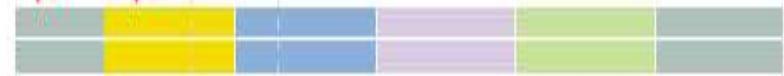
lacO
operator

X

X

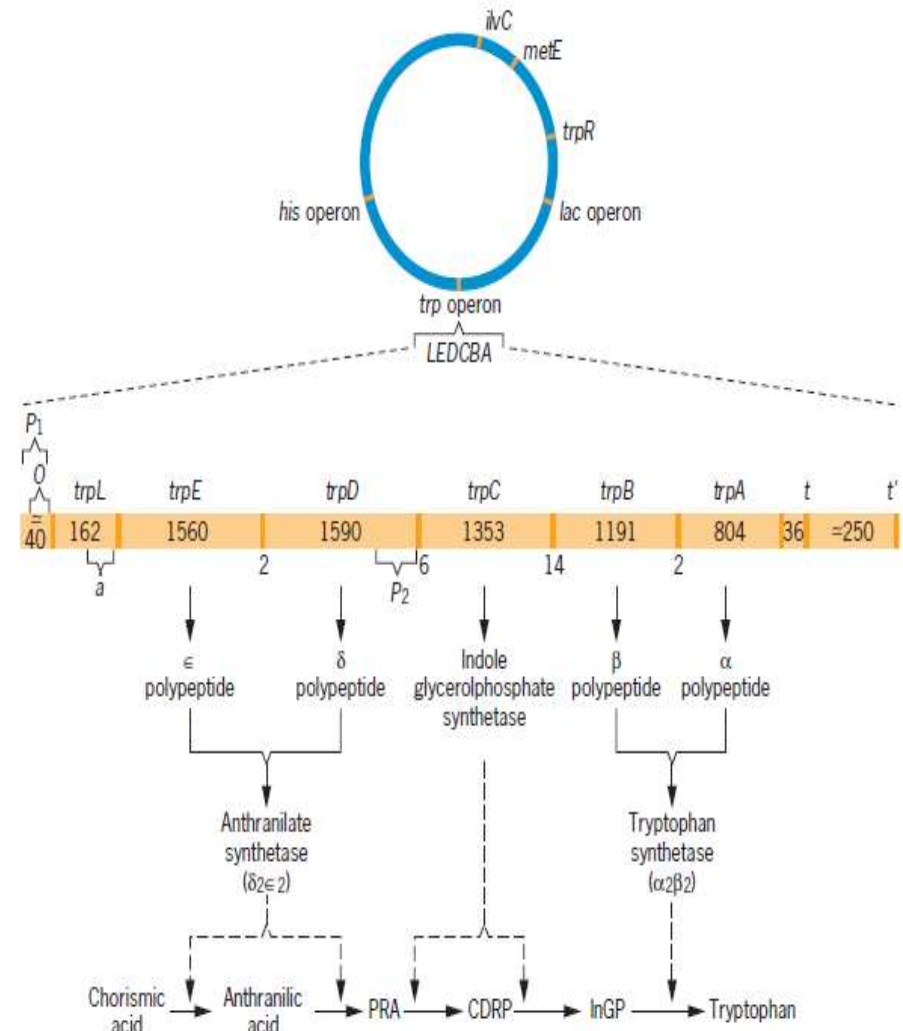
3 ...so transcription is at a low rate.

Little transcription



Tryptophan Operon system:

- This Operon is required for the synthesis of amino acid **tryptophan**
- It is an example for **Negatively Controlled Repressible Operon**
- A negatively controlled Operon- one that is inhibited by a regulatory protein such as a Repressor
- **Tryptophan Operon= Repressible Operon:** it is inhibited when an effector molecule is present
- Tryptophan Operon consists of:
 - 5 structural genes- ***trpE***, ***trpD***, ***trpC***, ***trpB*** and ***trpA***
 - These genes codes for enzymes involved in the conversion of **chorismic acid to tryptophan**
 - The controlling site in trp Operon lies next to ***trpE***
 - It consists of promoter site, ***trpP***
 - An overlapping operator site, ***trpO***
 - Leader region, ***trpL*** that codes for a leader peptide and an RNA attenuator
 - The leader peptide is involved in the attenuation mechanism

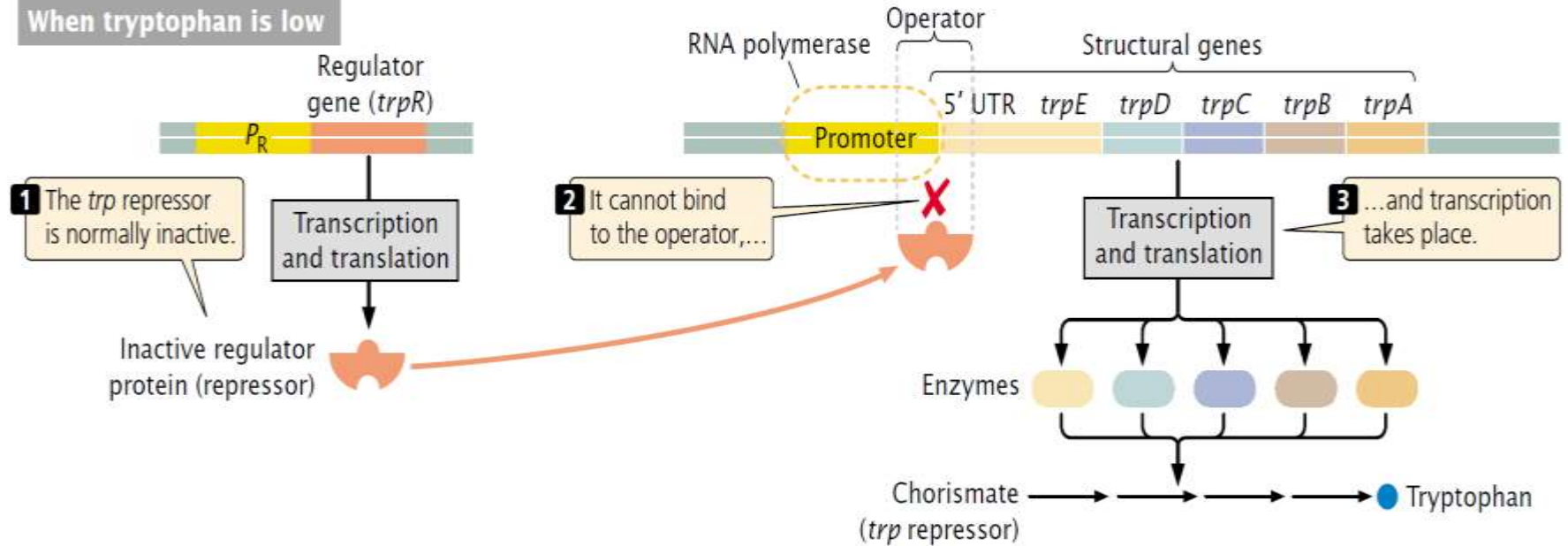


■ **FIGURE 18.13** Organization of the *trp* (tryptophan) operon in *E. coli*. The *trp* operon contains five structural genes that encode enzymes involved in the biosynthesis of tryptophan, as shown at the bottom, and the *trpL* regulatory region. The length of each gene or region is given in nucleotide pairs; the intergenic distances are shown below the gene sequence. Key: PRA, phosphoribosyl anthranilate; CDRP, carboxyphenylamino-deoxyribulose phosphate; InGP, indole-glycerol phosphate.

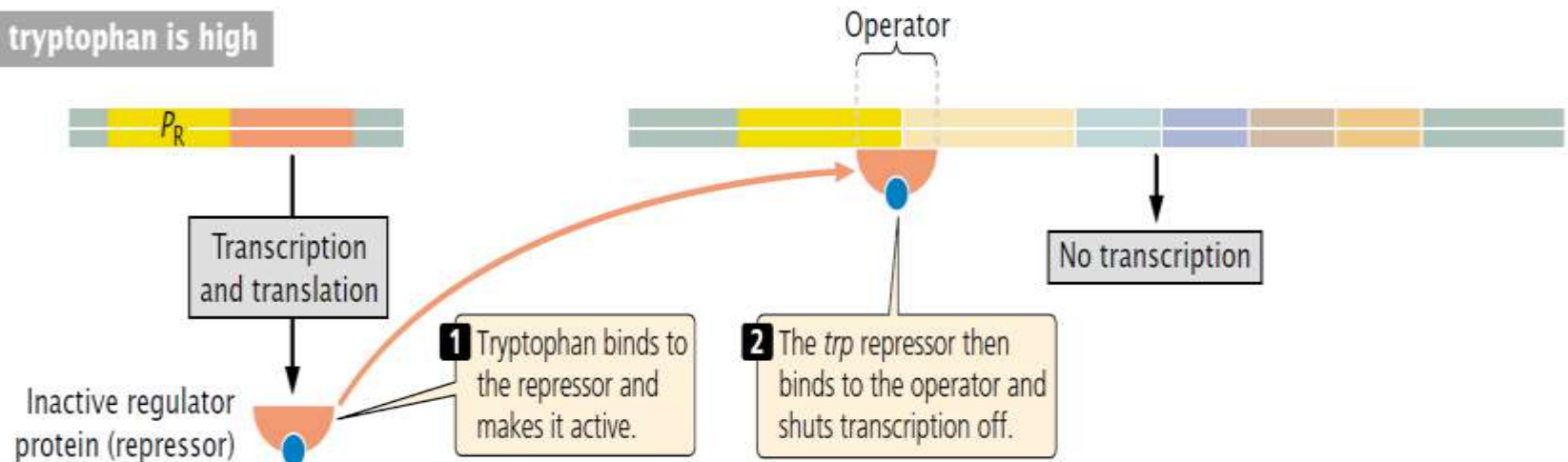
In this Repressible Operon,

- The product of the **trpR** gene, the repressor, is inactive itself (called **Aporepressor**)
- It does not recognize the operator sequence of the trp Operon
- The repressor becomes active only when, It combines with tryptophan
- In the **Absence of Tryptophan**,
 - Aporepressor (inactive regulatory protein) has no effect on the tryptophan Operon
 - The enzymes for synthesizing the amino acid are produced and the cell makes tryptophan
 - Tryptophan is thus called **Corepressor** since it acts as an **Effector** to convert the Aporepressor into the active repressor
 - When tryptophan binds to Aporepressor, it induces a conformational change that converts the regulatory protein into the tryptophan **Holorepressor**
 - The **Holorepressor** binds to **trpO** and blocks the attachment of **RNA polymerase** to **trpP**
 - This shuts **off the transcription** of the structural genes

When tryptophan is low



When tryptophan is high



When tryptophan is high

(a)

- 1 RNA polymerase begins transcribing DNA, producing region 1 of the 5' UTR.

(b)

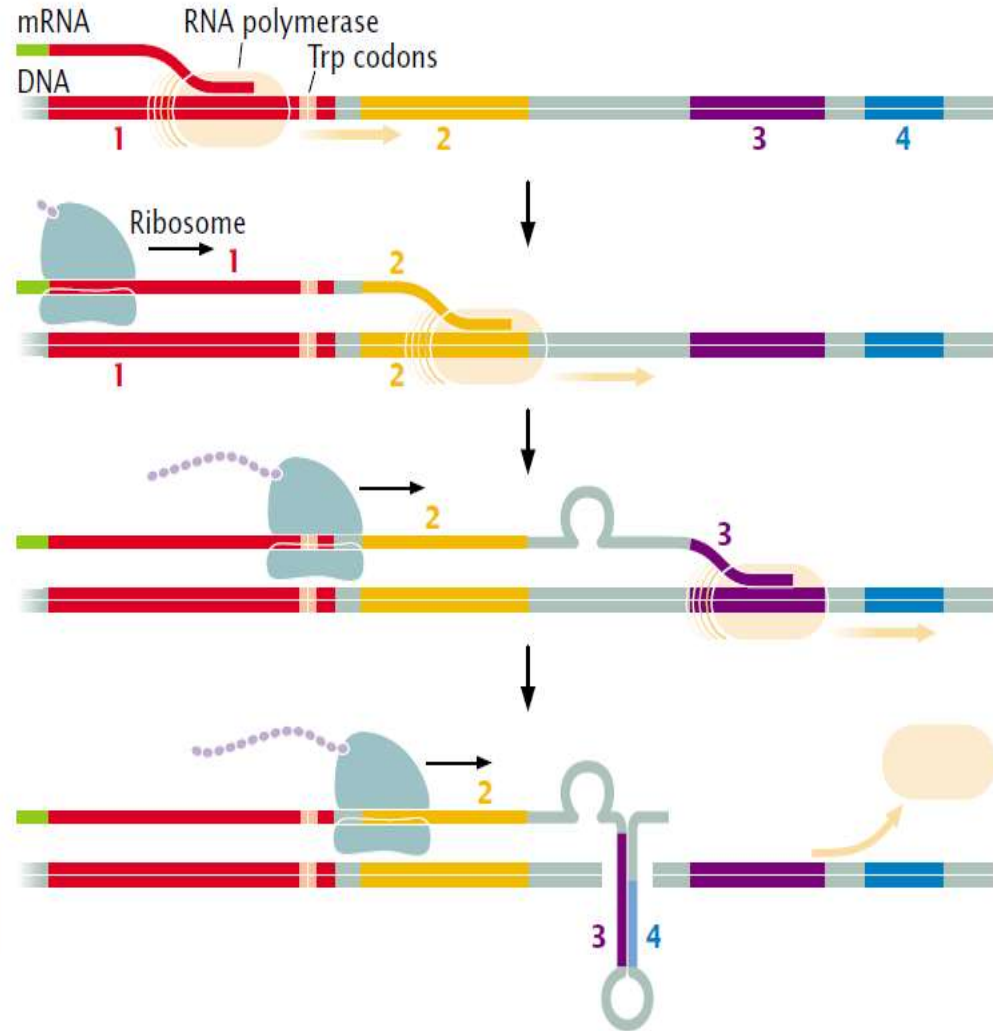
- 2 A ribosome binds to the 5' end of the 5' UTR and begins to translate region 1, while region 2 is being transcribed.

(c)

- 3 The ribosome translates region 1 while RNA polymerase transcribes region 3.
- 4 The ribosome does not stall at the Trp codons, because tryptophan is abundant.

(d)

- 5 The leading edge of the ribosome covers part of region 2, preventing it from pairing with region 3.
- 6 Region 4 is transcribed and pairs with region 3. The pairing of regions 3 and 4 produces the attenuator that terminates transcription.



When tryptophan is low

(e)

- 1 RNA polymerase begins transcribing the DNA, producing region 1 of the 5' UTR.

(f)

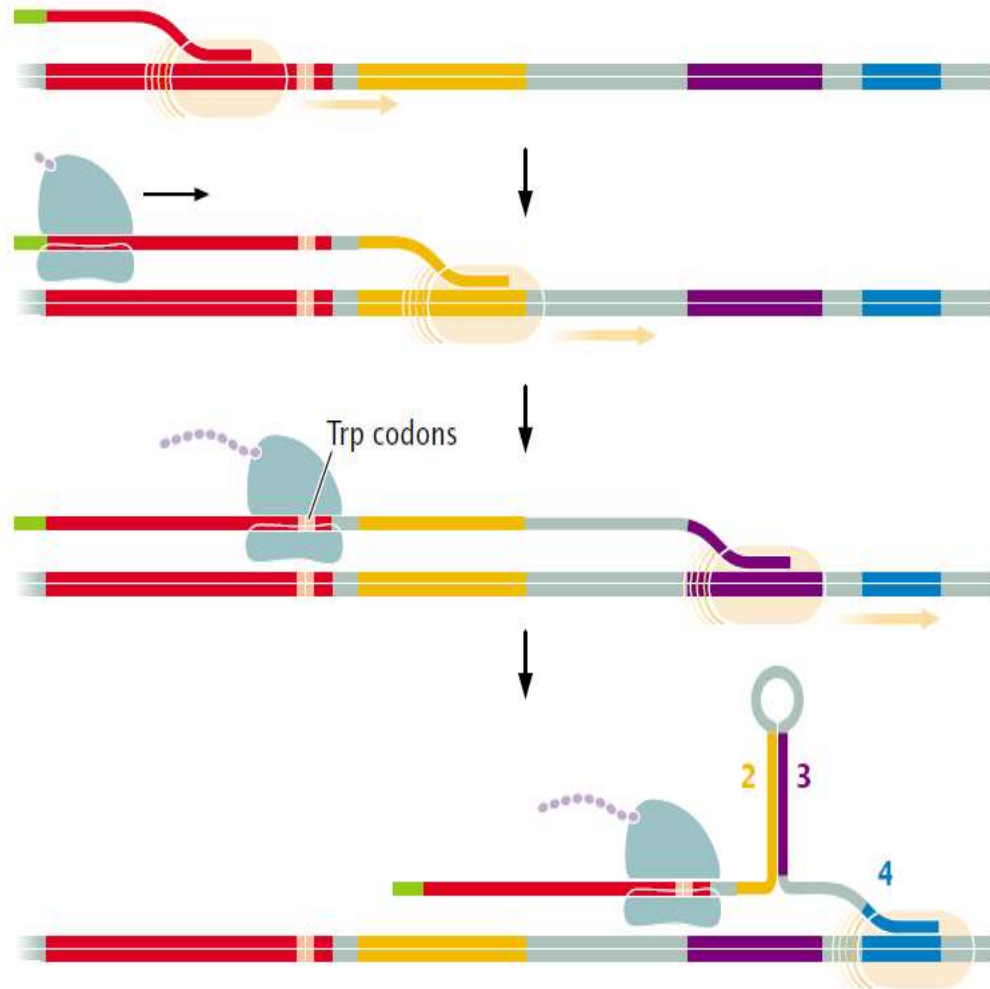
- 2 A ribosome attaches to the 5' end of the 5' UTR and begins to translate region 1 while region 2 is being transcribed.

(g)

- 3 The ribosome stalls at the Trp codons in region 1 because tryptophan is low.
- 4 Because the ribosome is stalled, region 2 is not covered by the ribosome when region 3 is transcribed.

(h)

- 5 When region 3 is transcribed, it pairs with region 2.
- 6 When region 4 is transcribed, it cannot pair with region 3, because region 3 is already paired with region 2; the attenuator never forms, and transcription continues.

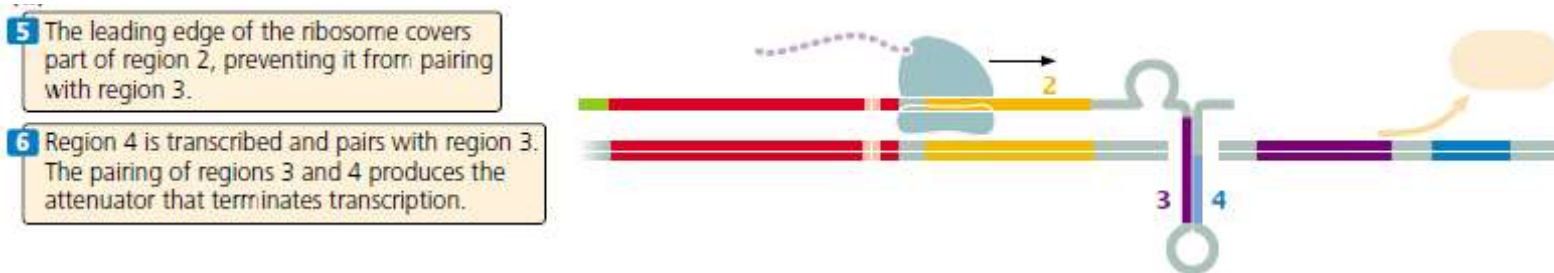


Transcriptional attenuation:

- It is a mechanism which negatively regulates the expression of many operons in bacteria
- It controls the ability of RNA polymerase to continue elongation past specific sites
- Transcriptional attenuation is carried out by 2 different processes:
 1. Translation-mediated
 2. RNA-binding proteins mediated

1. Translation mediated transcriptional attenuation:

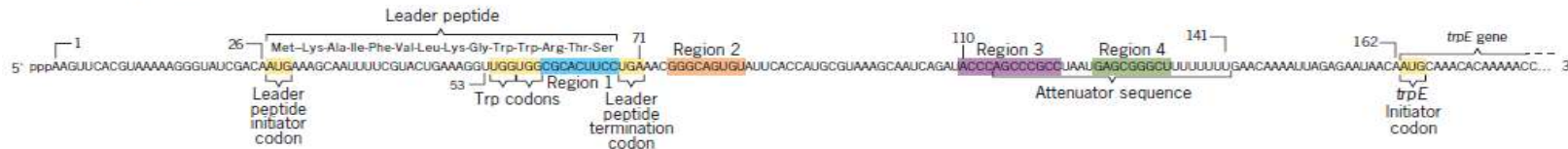
- It provides an additional level of control in *E coli*
- In trp Operon, the site of attenuation is located downstream of the transcription start site
- This mechanism depends on the fact that translation in bacteria is coupled with transcription, so ribosomes begin translating the 5' end of an mRNA while it is still being synthesized
- **If Tryptophan is abundant, most transcription terminates at this site**



- **If Tryptophan is scarce, transcription continues to yield functional trp mRNA**



Regulatory components of the *trpL* region



(a)

Alternate secondary structures formed by the *trpL* transcript

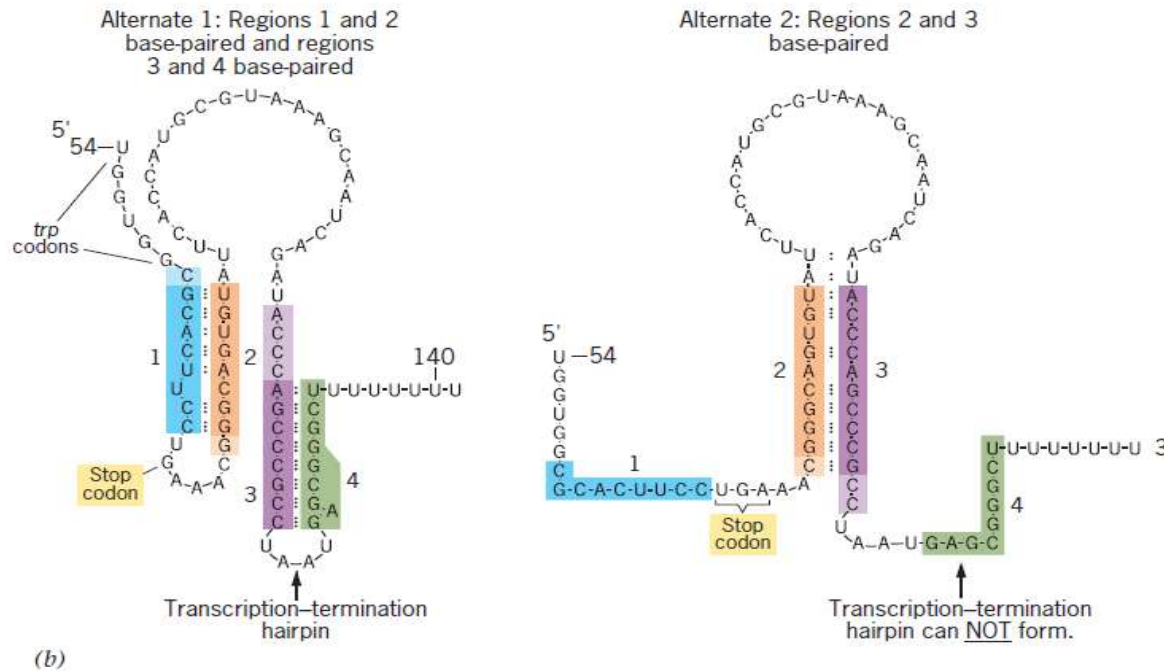


FIGURE 18.14 Sequences in the leader region of the *trp* mRNA responsible for attenuation. (a) The *trpL* sequence, highlighting the sequence encoding the leader peptide, the two tandem tryptophan codons responsible for the control of attenuation by tryptophan, and the four regions (shaded) that form the stem-and-loop or hairpin structures shown in (b). (b) Alternate secondary structures formed by the *trpL* mRNA—either (1) region 1 will pair with region 2 and region 3 with region 4, forming a transcription-termination hairpin, or (2) region 2 will base-pair with region 3, preventing region 3 from pairing with region 4. The concentration of tryptophan in the cell determines which of these structures will form during the transcription of the *trp* operon.

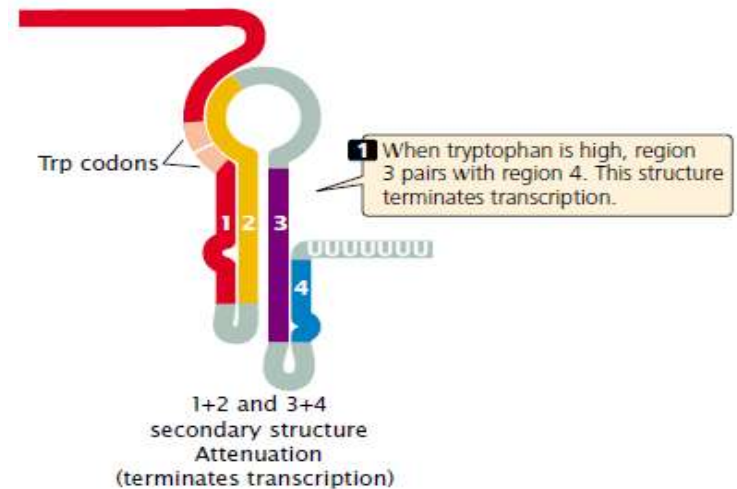
The **leader sequence** (162 bp) of the Operon has 14 codon open reading frame that includes,

2 codons of tryptophan
This short coding sequence codes for a leader peptide of 14 amino acids

The ribosome attempts to translate the leader transcript

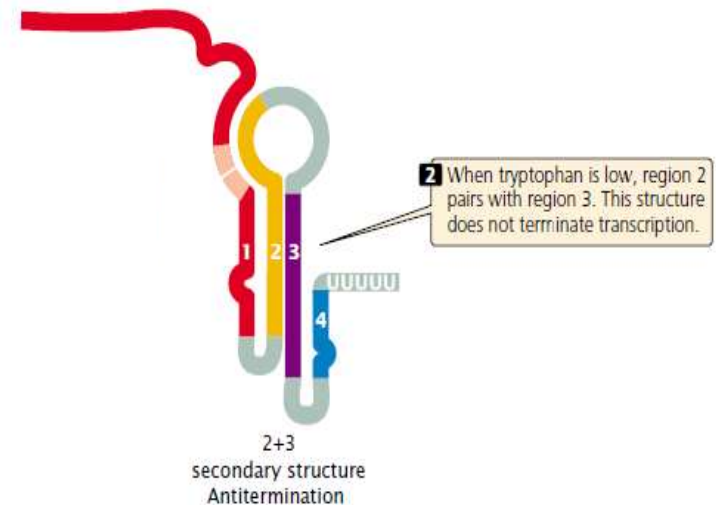
Under excess tryptophan,

- The termination of transcription is signaled by a **stem-loop structure** that forms by complementary base pairing between 2 specific sequences of the growing trp mRNA chain
- The structure forms if the translation of the growing chain is proceeding at a normal rate
- The **3-4 stem-loop forms** (the terminator stem), terminating transcription and translation continues



• Under low tryptophan,

- Protein synthesis stalls at a critical region of the message
- The ribosomes bound to the mRNA block formation of the transcription-terminating stem loop, allowing trp mRNA synthesis to continue
- The ribosome is stalled, and the **stem-loop 2-3 forms**, allowing continued transcription



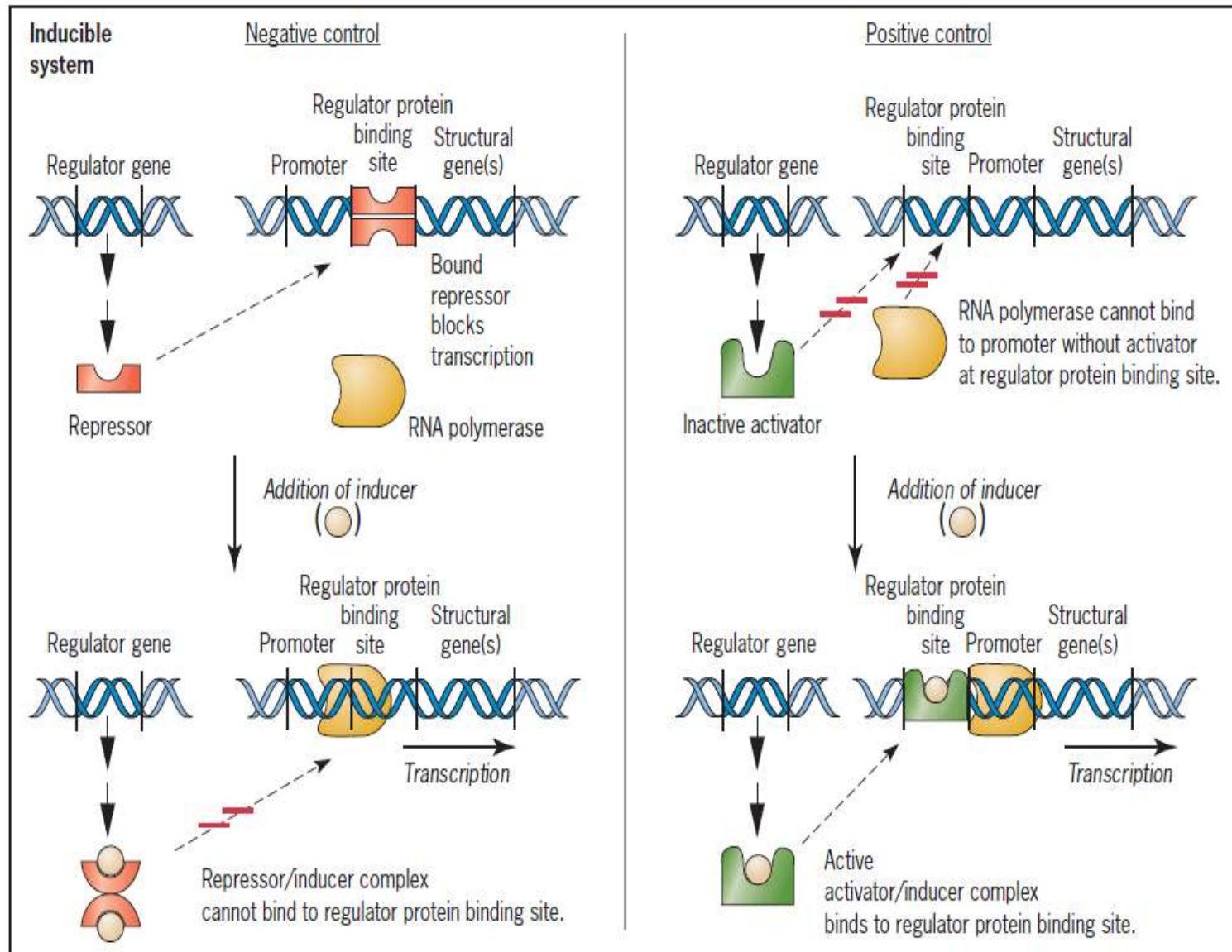
The Critical Region of trp mRNA contains 2 adjacent tryptophan codons,

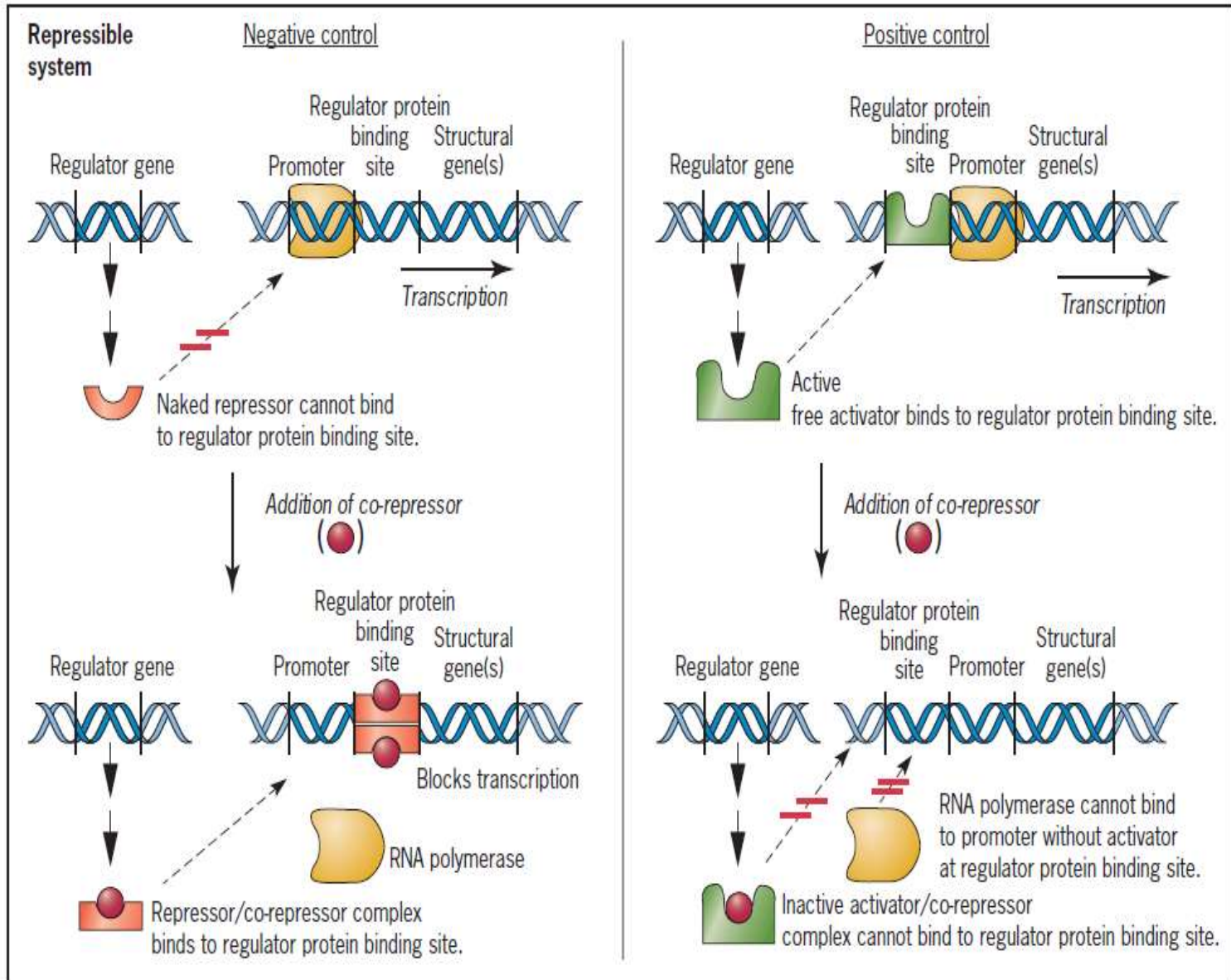
- **If tryptophan level in the cell is low= the ribosome stalls at this point and transcription of trp mRNA continues**
- **If tryptophan is abundant, translation continues and transcription is terminated**

2. RNA binding protein mediated attenuation:

- This mechanism happens in *B subtilis*
- RNA –binding terminator protein called **TRAP (Trp RNA binding Attenuation Protein)** mediates the transcriptional attenuation of trp Operon
- TRAP forms a multimer of 11 subunits and each subunit binds a single tryptophan amino acid
- TRAP is activated by binding of tryptophan
- TRAP binds to a sequence in the leader segment of transcript
- Formation of terminator hairpin
- The termination of transcription then prevents the production of the tryptophan biosynthetic enzymes
- In the **Absence of Tryptophan**,
 - TRAP does not bind
 - mRNA adopts a structure that prevents the terminator hairpin from forming
 - Protein **anti-TRAP** controls the activity of TRAP, which binds to TRAP, and prevents it from repressing the tryptophan Operon

An overview of inducible and repressible Operon under negative and positive control





Riboswitches:

- Cis acting regulatory elements
- mRNA elements that bind **metabolites or metal ions** as ligands and regulate mRNA expression by forming alternative structures in response to this ligand-binding
- Ex. **Magnesium ions, nucleic acid precursors, enzyme cofactors** and **amino acid residues**
- Riboswitches are often located in the **5' UTR** of bacterial mRNA
- Function of Riboswitches is associated with the ability of RNA to form a diversity of structures
- These are composed of 2 domains:
 - **Aptamer domain:** Acts as receptor that specifically binds a ligand
 - **Expression platform:** Acts directly on gene expression in response to ligand binding
- Changes in the conformation of Riboswitches happen after binding of ligands with the Aptamer domain
- This in turn controls gene expression by one or two mechanisms,
 - **Premature termination of transcription (i.e. attenuation)**- the **Riboswitch** controls whether or not the mRNA for the biosynthetic genes will be prematurely terminated
 - **Translational inhibition**- the Riboswitch controls whether or not the mRNA will be translated

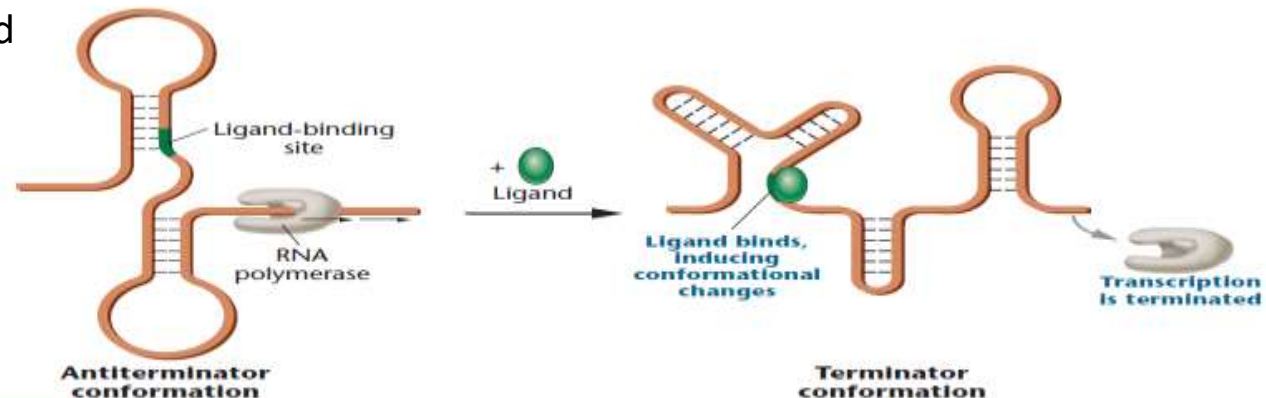
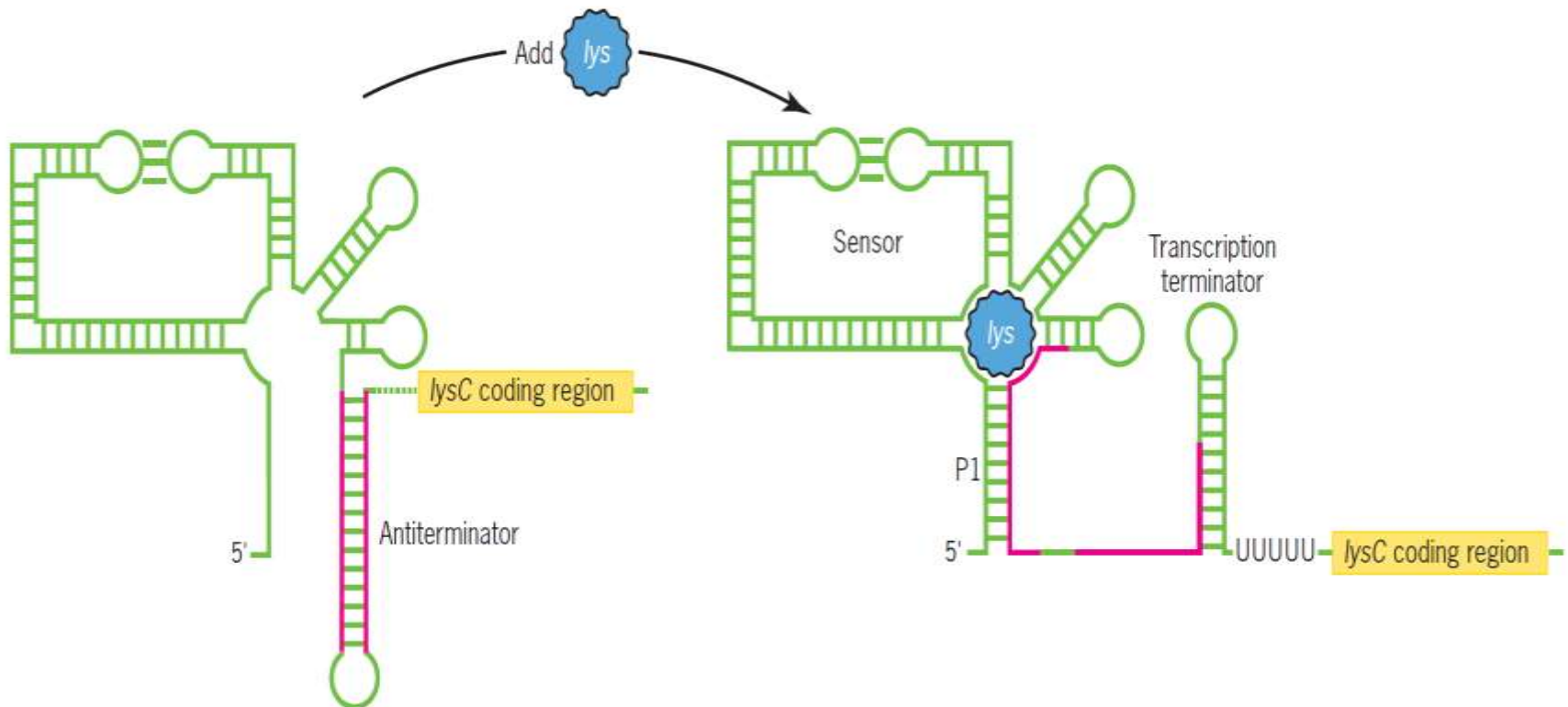


FIGURE 16-15

Illustration of the mechanism of riboswitch regulation of gene expression, where the default position (left) is in the antiterminator conformation and, upon binding by the ligand, adopts the terminator conformation.

An example for Riboswitches- Lysine Riboswitch



(a) Lysine absent, *lysC* gene is transcribed.

(b) Lysine present, *lysC* transcription is terminated.

■ **FIGURE 2** Control of transcription by the lysine riboswitch. (a) In the absence of lysine, an antiterminator hairpin forms in the expression domain upstream from the start of translation of the regulated gene (*lysC* in the diagram) and precludes the formation of the transcription-terminator hairpin. As a result, transcription of the gene is turned on. (b) When present, lysine is bound by the aptamer, and the upstream antiterminator sequence is sequestered in a basal hairpin that is part of the lysine-binding pocket. As a result, the downstream antiterminator sequence is free to participate in the formation of a transcription-termination hairpin upstream from the coding region of the regulated gene (*lysC* is shown). Therefore, transcription is turned off.

Bacteriophage lambda: a transcriptional switch

- Bacteriophage lambda
- Temperate phage
- Infects the bacterium *E coli*
- Its genome = linear dsDNA, 48.5 kbp, contains about 50 genes
- After infecting host cells (*E coli*) lambda phage can follow one of two alternative cycles:
 1. **Lysogenic cycle**
 2. **Lytic cycle**

1. Lysogenic cycle:

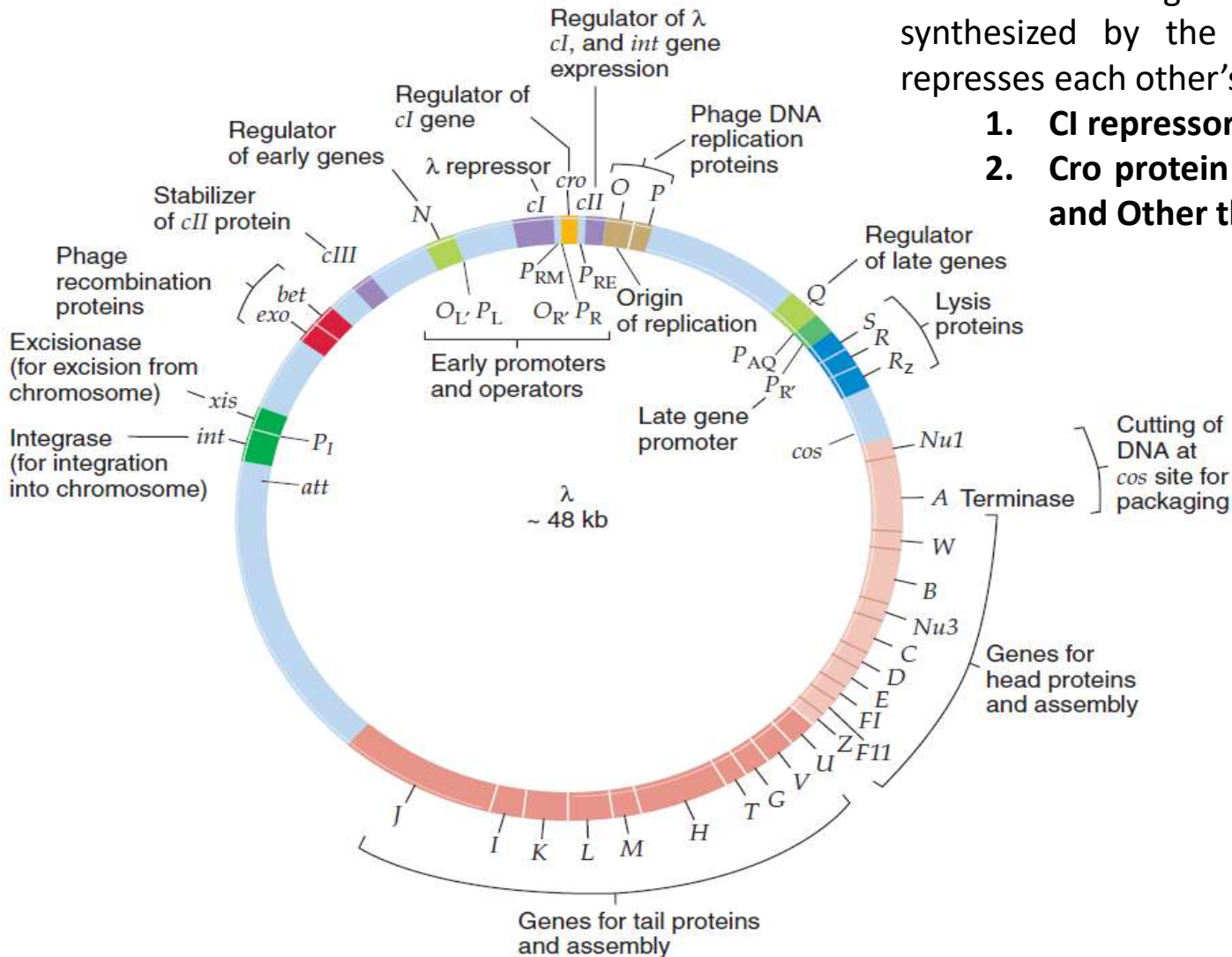
- Phage DNA integrates with the bacterial genome
- In this stage, phage genome is present in the bacterium in a latent form known as **Prophage**
- Lambda phage has an Operon-type system controlling its two functional states
- The switch between these two states is mediated by proteins encoded by the bacteriophage genome

2. Lytic cycle:

- The phage DNA enters the host bacterium
- Its genes are transcribed
- The phage genetic materials replicated
- Protein components of phage particles are produced
- The host bacterium is broken open (lysed) to release the assembled progeny particles by the process called lysis

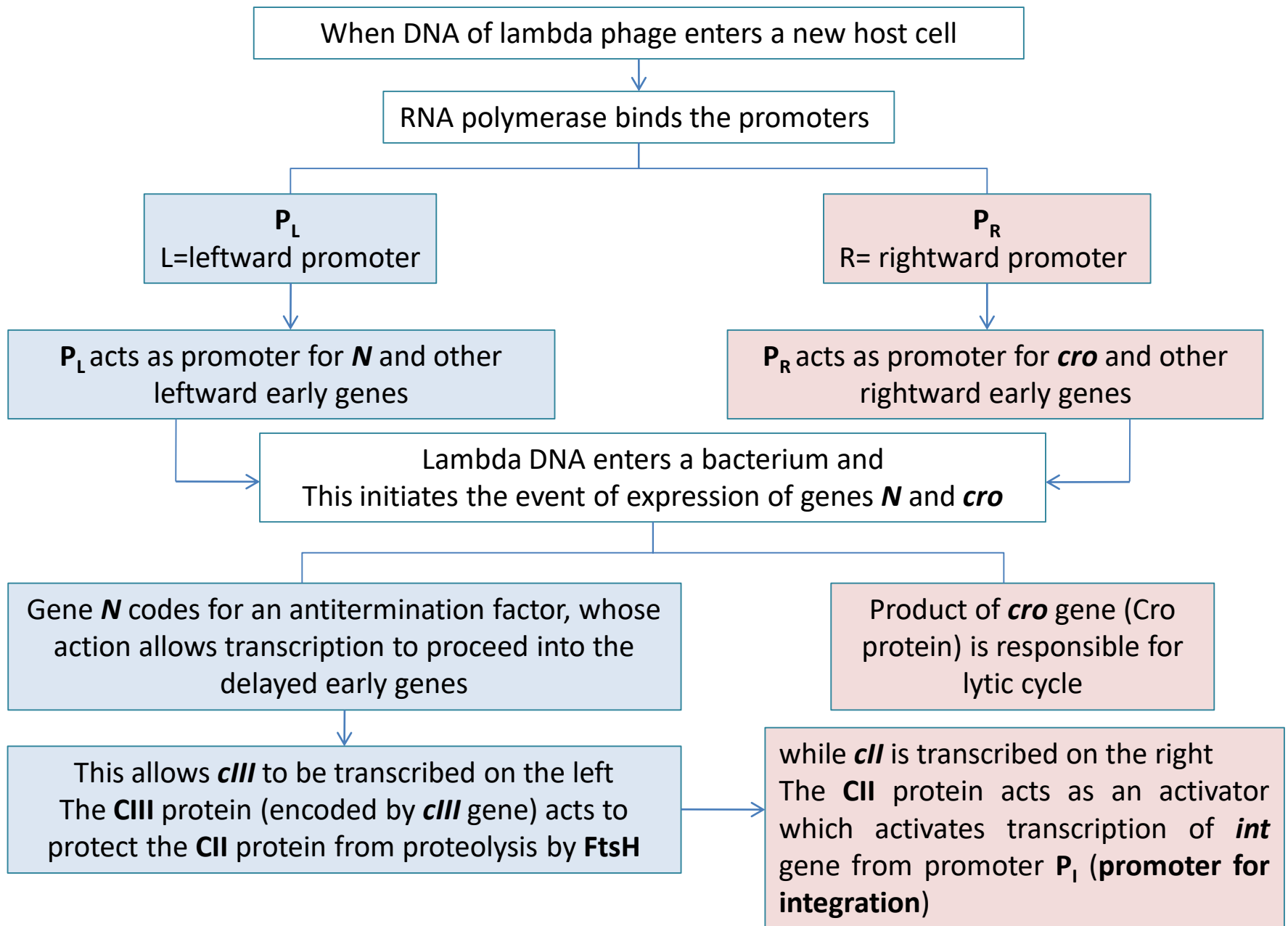
Figure 17.21

A map of phage λ , showing the major genes. (Promoters discussed in text: P_L = promoter for leftward transcription of the left early operon, P_R = promoter for rightward transcription of the right early operon, P_{RE} = promoter for repressor establishment, and P_{RM} = promoter for repressor maintenance.)



There are 2 gene regulatory proteins synthesized by the virus; These proteins represses each other's synthesis .

1. **CI repressor (Lambda repressor)**
2. **Cro protein (Control for Repressor and Other thing)**



The *cro* and *cII* genes are adjacent to each other and between them there is another promoter-called P_{RE} (RE Repressor Establishment)
This promoter can be recognized by RNA polymerase only in the presence of **CII**
Transcription from P_{RE} results in expression of ***cl***

The two promoters (P_{RM} and P_R) can never be occupied simultaneously
Thus ***cl*** and ***cro*** genes are never transcribed at the same time

The genes ***cl*** and ***cro*** are adjacent
Between them there is a promoter region of ***cl*** gene – P_{RM} (RM- Repressor Maintenance)
The ***cl*** gene is also transcribed from promoter P_{RM}

If RNA polymerase binds to P_{RM} promoter,
CI repressor mRNA will be produced

Gene ***cl*** encodes **CI repressor (lambda repressor)**
which is responsible for **lysogenic cycle**

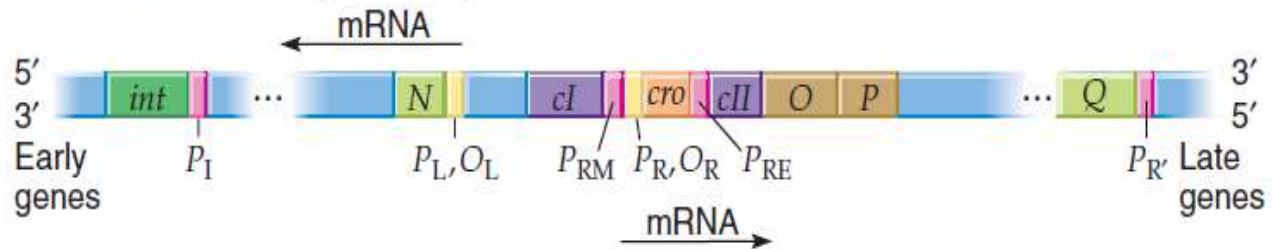
If RNA polymerase binds the P_R promoter,
cro mRNA will be produced

Product of ***cro*** gene (Cro protein) is
responsible for **lytic cycle**

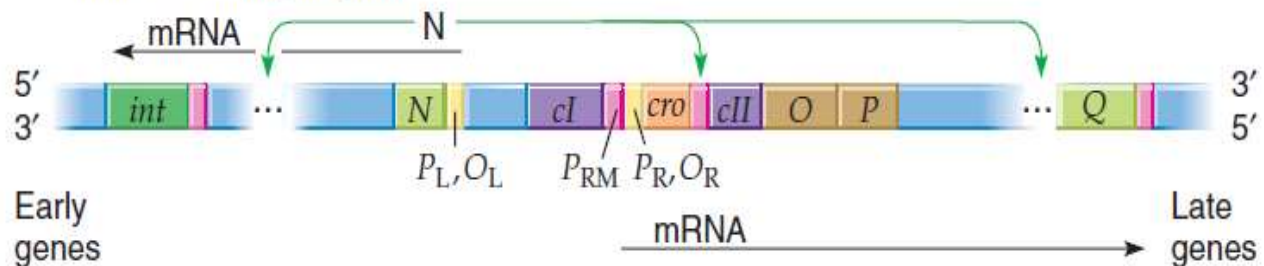
Figure 17.22

Expression of λ genes after infection of *E. coli*, and the transcriptional events that occur when either the lysogenic or lytic pathway is followed. In the figure, stimulation of transcription is indicated by green arrows, and repression of transcription by red arrows.

- 1** Phage growth begins when RNA polymerase binds early promoters P_L and P_R , making mRNA for *N* and *cro* genes. (Throughout figure, are inactive promoters and are active promoters.)



- 2** N protein, acting at three sites, extends RNA synthesis to other genes; O and P proteins permit DNA synthesis to begin; cII protein is made and acts as shown in the following diagram:

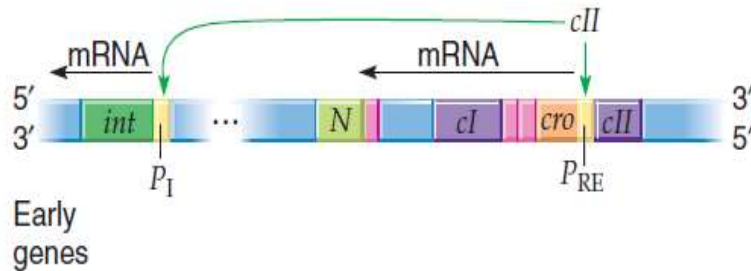


3 Two competing pathways occur:

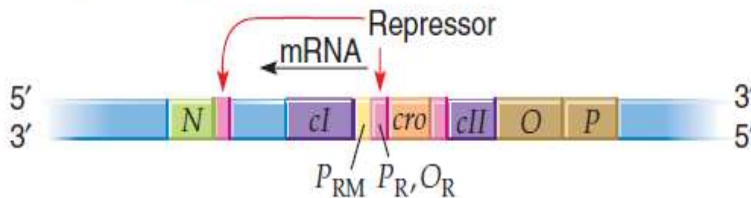
Toward lysogenic development:
dominance of repressor (*cI*)

Toward lytic development:
dominance of Cro

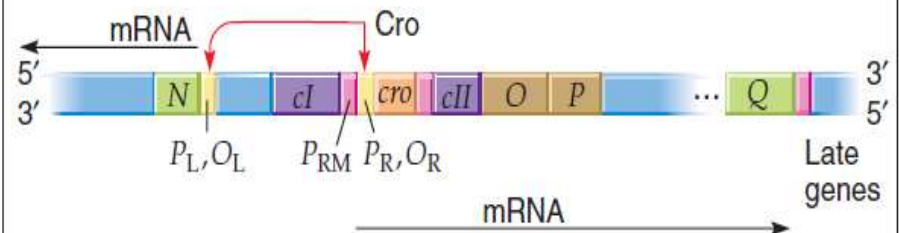
- 4a** *cI* protein stimulates synthesis of mRNA from P_{RE} for repressor (*cI* gene) and from P_I for integrase (*int* gene).



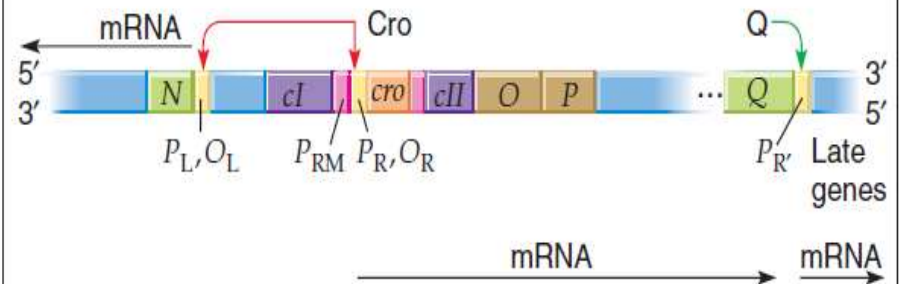
- 5a** Sufficient repressor binds O_R and O_L , blocking RNA synthesis; repressor bound to O_R stimulates synthesis of more repressor from P_{RM} . Integrase promotes integration of phage DNA. Lysogeny is established.



- 4b** Cro protein occupies operators O_R and O_L , preventing synthesis of repressor mRNA, but allowing enough rightward transcription for Q protein to accumulate.



- 5b** Q protein stimulates late gene transcription; these genes encode structural proteins that package DNA into new phage particles.



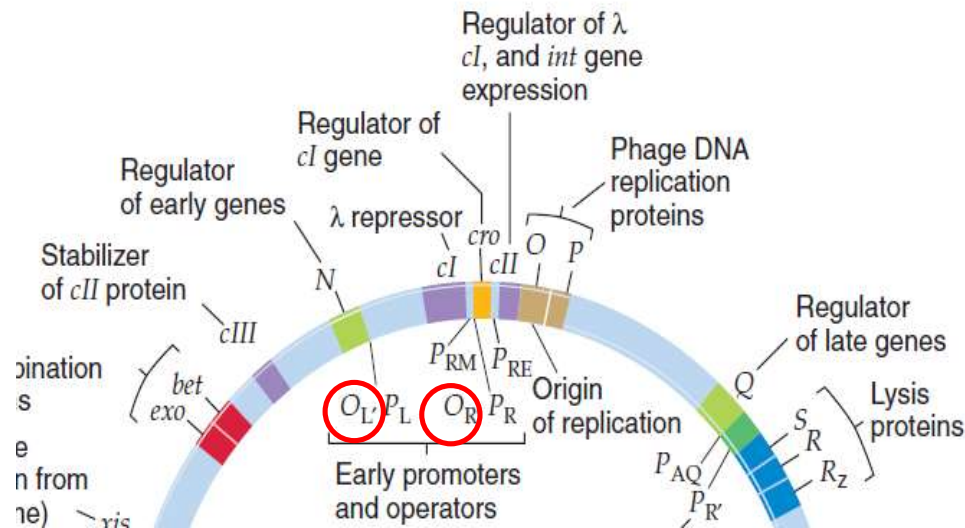
Promoter P_L and P_R are associated with an operator O_L and O_R respectively at which repressor protein binds to prevent RNA polymerase from initiating transcription

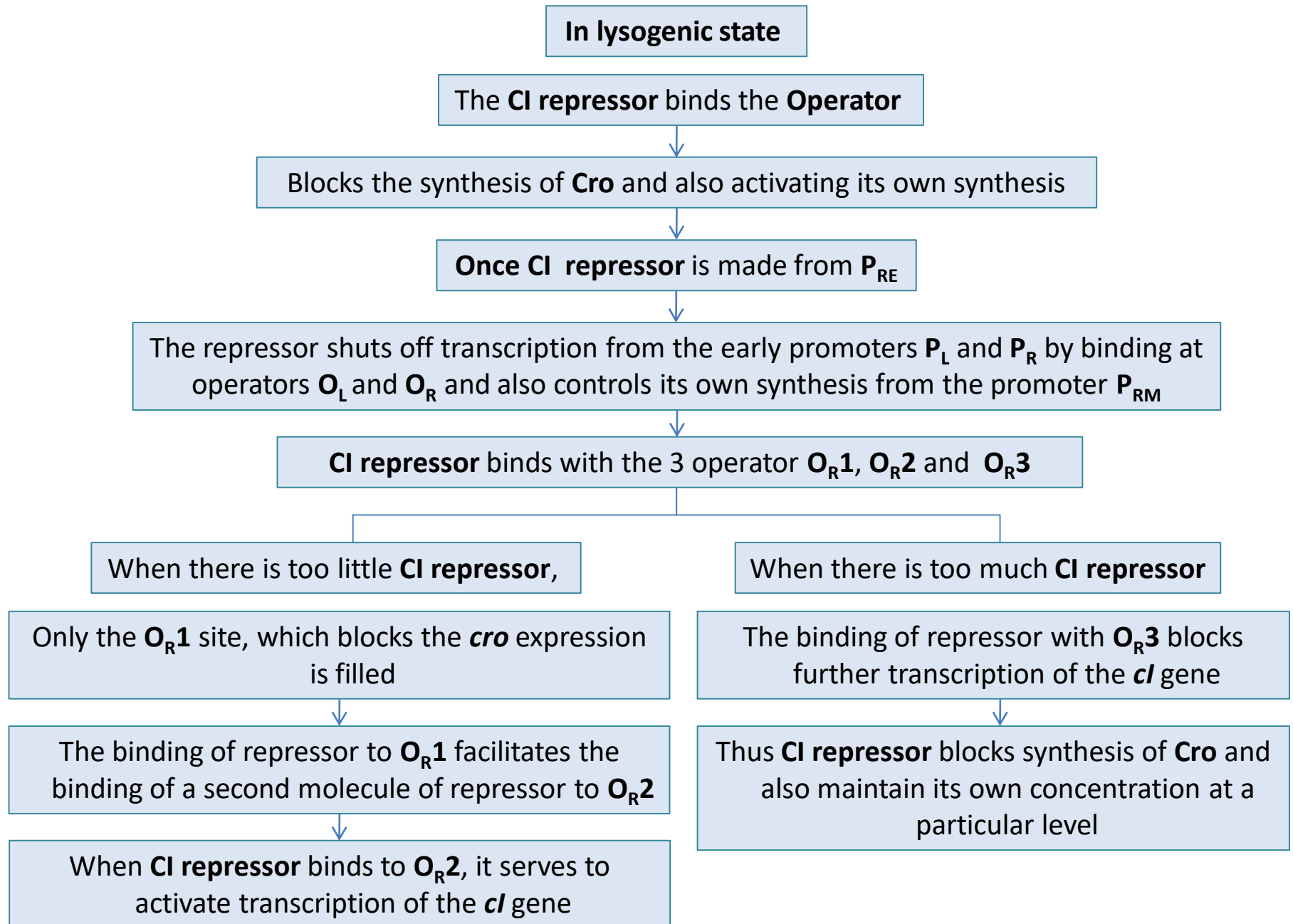
Each operator contains 3 Repressor-binding Sites:

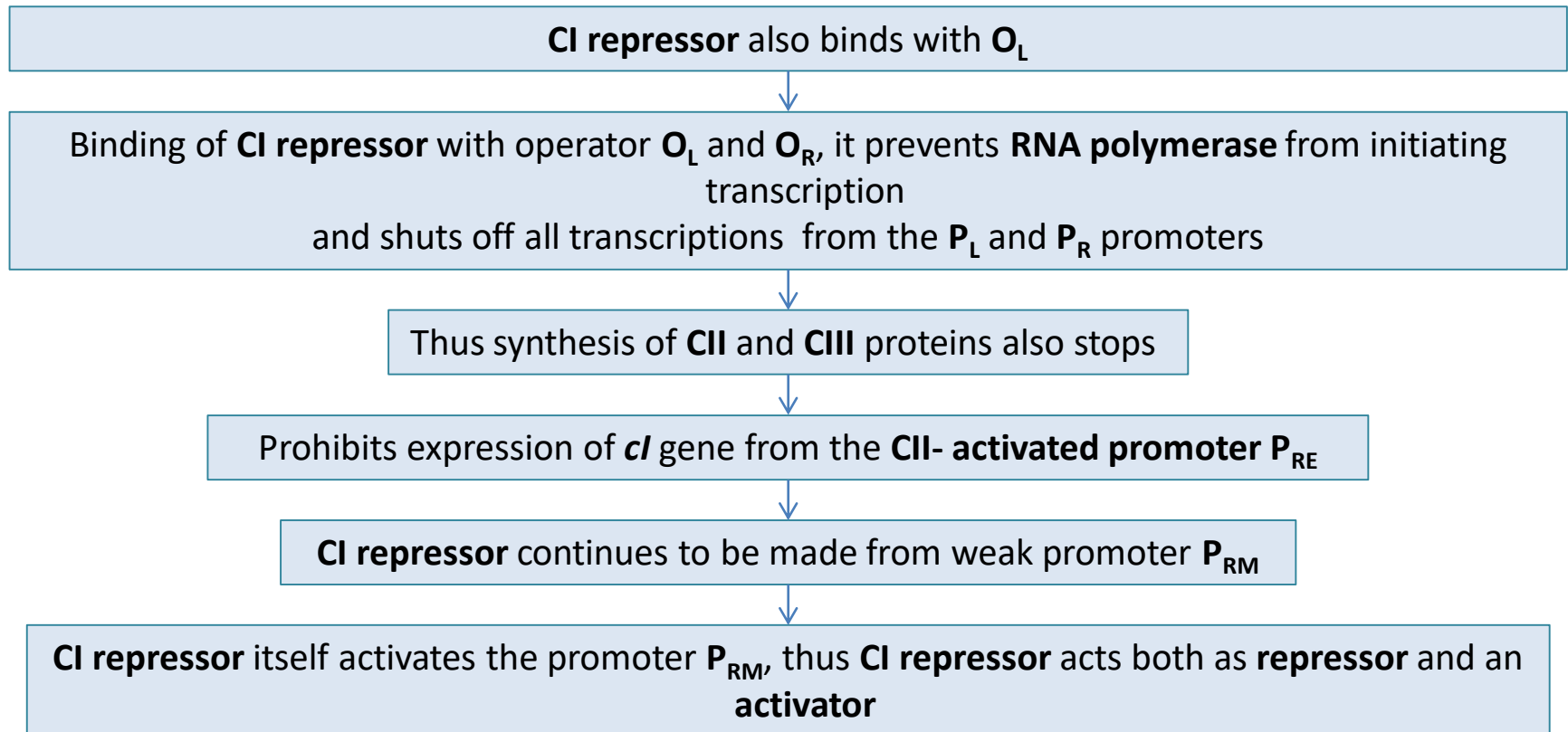
O_L consists of series binding sites- $O_L1-O_L2-O_L3$

O_R consists of series of binding sites- $O_R1-O_R2-O_R3$

In each case, **site 1** lies closest to the transcription start site in the promoter, and **sites 2 and 3** lie farther upstream







Int proteins is also required for lysogeny

Int allows the integration of the phage genome into the bacterial chromosome

CII activates transcription of *int* gene from promoter **P_I** (**promoter for integration**)

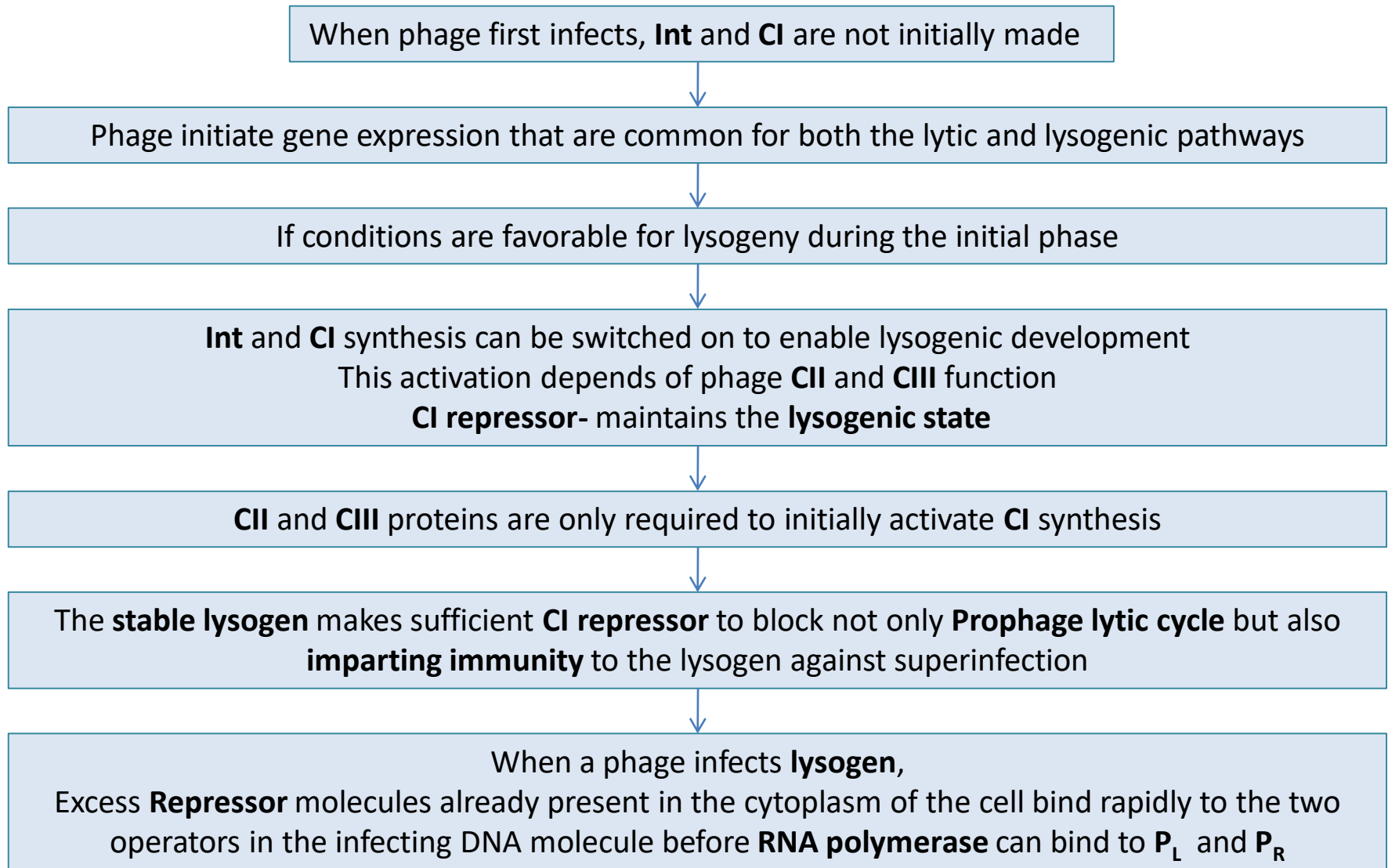
CII plays most important role in establishing lysogeny and it is a very unstable protein

The levels of **CII** are reduced by bacterial **FtsH** (**HflB**) protease, encoded by *hfl* (high frequency of lysogeny) gene

CIII is also important to establish lysogeny (in indirect way)

CIII inhibits the **FtsH protease** activity

Thus in the absence of **CIII**, **FtsH protease** degrades **CII** and not enough **CII** is available to establish lysogeny



Retroregulation of *int* gene expression:

- **CII** is required for **Int** expression following lambda infection
- *int* gene transcription is activated from the **P_I** promoter by **CII**
- *sib* segment inhibits *int* expression from **P_L** but not from **P_I**
- The sequence of *sib* site allows the formation of a large **stem-loop structure** in the RNA, which is sensitive to cleavage by **RNaseIII**
- Thus cleavage, of the transcript initiated from **P_L promoter** (**P_L transcript**) at *sib* by **RNase III** prevents **Int** protein synthesis
- Thus *sib* is required to retroregulate *int* expression from a downstream site
- Retro= because *sib* is located downstream from *int*

In lytic state,

cro gene is expressed from the P_R promoter

It prevents the expression of *ci* gene

Cro forms a small dimer that prevents the expression of lambda repressor from P_{RM}

It is a weak repressor that also binds with both O_R and O_L

Binding of O_L first with O_{R3}

Then **Cro** binds to O_{R2} or O_{R1}

This inhibits **RNA polymerase** from initiating transcription from P_{RM}

So **Cro's** first action is to prevent the lysogeny

Cro has 2 responsibilities:

1. To prevent the synthesis of CI repressor
2. Turning down the expression of the early genes once there has been enough product made

Product of **Q** gene is required for **Capsid Formation** and **Lysis**

The **Q protein** activates transcription initiation at P_R (secondary rightward promoter)
It acts as an antiterminator protein allowing transcripts initiated at P_R to proceed through the termination signal t_R

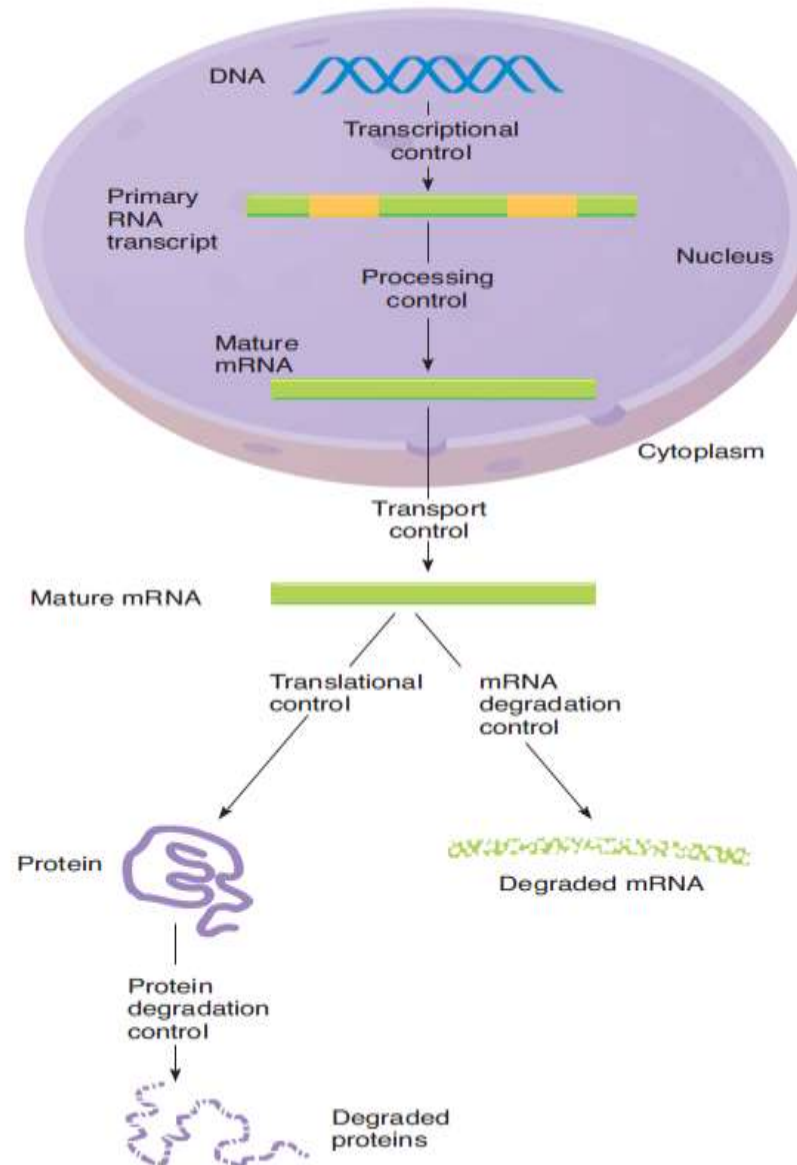
Lysogenic cycle:	Lytic cycle:
Bacterial cell undergoes lysogenic cycle when conditions are poor for bacterial growth	Bacterial cell undergoes lytic cycle when the phage infects a population of bacterial cells that are growing vigorously
<i>E coli</i> lacking FtsH always forms lysogen as it allows CII to synthesize sufficient CI repressor	<i>E coli</i> having FtsH which degrades CII (CII responsible for CI repressor synthesis)
If growth condition is poor = no FtsH = CII not destroyed= CI repressor is made sufficiently	If growth condition is good= FtsH is active= CII is destroyed efficiently= CI repressor is not made
CI repressor synthesized more quickly	Cro repressor synthesized more quickly
<i>cl</i> expression happens and silences the <i>cro</i> gene	Cro repressor blocks <i>cl</i> expression and <i>cro</i> genes expresses

Regulation of transcription in eukaryotes:

- More complex
- Regulation of gene expression happens in 2 levels,
 - At transcription initiation level
 - At post transcriptional level (happen during RNA processing, RNA transport (nucleus to cytosol), translation and post translational modification of a protein)
- Transcription is controlled by proteins that bind to specific regulatory sequences and modulate the activity of RNA polymerase
- Packaging of DNA into chromatin
- Chromatin modification

Figure 18.1

Levels at which gene expression can be controlled in eukaryotes.



Influence of chromatin structure on transcription:

Changes in chromatin structure= Repression or Stimulation of gene expression

As genes become more transcriptionally active

Chromatin shows increased sensitivity to **DNase I** digestion

Increased sensitivity to DNase I suggests
chromatin structure is more open
more relaxed or open conformational state for chromatin, allowing access to
transcriptional factors as well as DNase I

DNase I hypersensitivity sites are found near regulatory promoter elements and also
occasionally at 3' end of the gene as well

Local alterations in chromatin structure facilitate the assembly of the transcription factors
and the RNA polymerase holoenzyme at the promoter

2 most important ways for altering chromatin structure are:

Covalent histone modifications:

- N terminal tails of histones undergoes post-translational covalent modifications
- 4 main types of modifications to histone tails regulates the gene expression:

1. **Acetylation of lysine**
2. **Methylation of lysine and arginine**
3. **Ubiquitinylation of lysine**
4. **Phosphorylation of serine and threonine**

Other less common modifications are:

1. **ADP- Ribosylation of glutamic acid**
2. **Sumoylation of lysine residues**

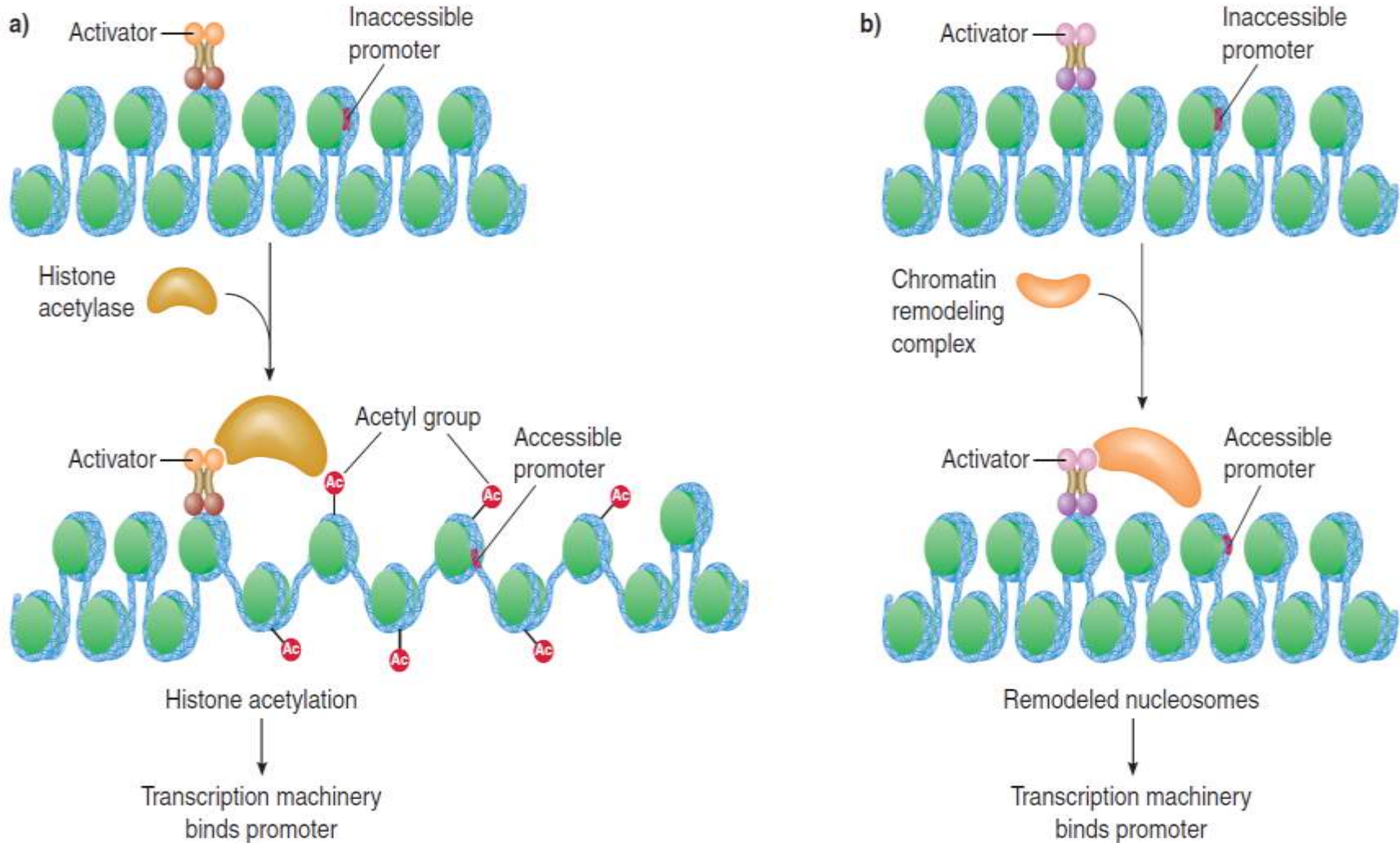
Nucleosome remodeling:

- Type of chromatin modification that influences the genome expression
- Involves a change in the structure of the nucleosome but no change in its position
- In some cases, the remodeling can involve cis and trans displacement that involves movement of nucleosome
- Induced by an energy dependent process that weakens the contact between the nucleosome and the DNA
- ATP- dependent chromatin remodeling complexes responsible for nucleosome remodeling
- 2 classes of complexes:
 - **Swi/snf family**
 - **ISWI-based family**

Figure 18.10

Chromatin modeling by (a) histone acetylases and (b) nucleosome remodeling complexes.

In either case, the result is access to the promoter by the transcription machinery.



1. Histone acetylation:

Lysines in the tails of histones are acetylated in chromosomal regions containing genes that are active or potentially active

Enzyme- **Histone Acetyl Transferases (HATs)** – functions as **Transcriptional Co-activators**

Histone deacetylation:

Removes acetyl group and promotes transcriptional repression

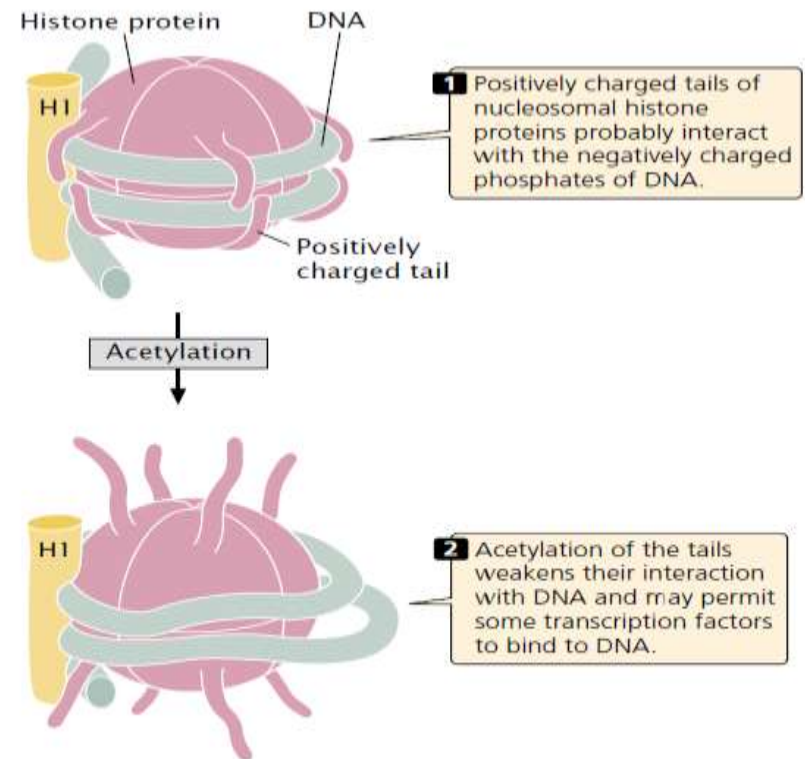
Enzyme- **Histone Deacetylases (HDACs)**- functions as **Transcriptional Repressors**

When charged lysine side chains are modified

Loss of positive charges

Affinity between histones and DNA reduces

RNA polymerase and transcription factors find it easier to access the promoter region



16.20 The acetylation of histone proteins alters chromatin structure and permits some transcription factors to bind to DNA.

Methylation:

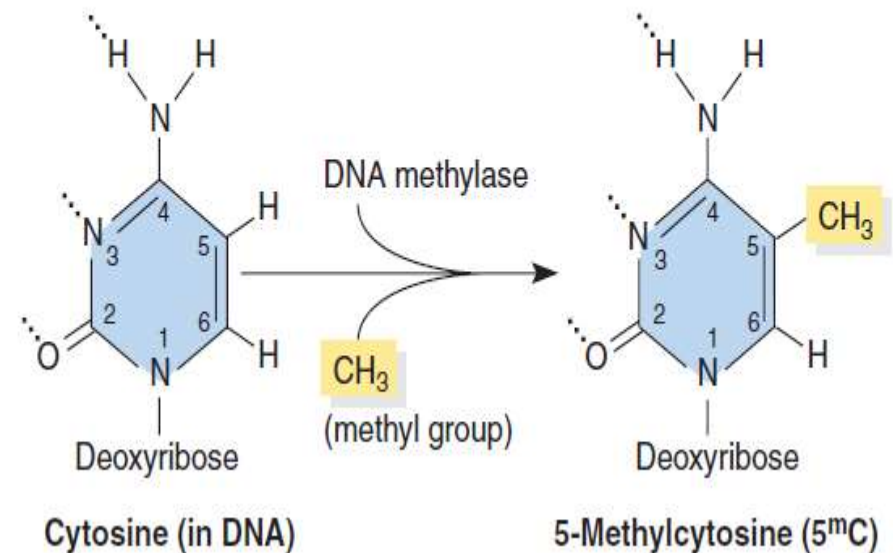
- Reversibly modifies histones
- Enzyme- **methyl transferases**- targets specific Arginine and Lysine residues in histones H3 and H4
- Histone methylation acts as a signal for both transcriptional repression and activation
- Methylation of histone H3 at lysine 9 (**H3K9**) and Histone H3 at lysine 27 (**H3K27**) –associated with gene silencing and heterochromatin formation
- Methylation of **H3K4, H3K36, H3K79**- associated with active chromatin
- Main function of DNA methylation:
 - In bacteria, both adenine and cytosine can be methylated
 - It provides protection from own restriction endonucleases and it also plays important role in mismatch repair
- 2 types of methylation processes in eukaryotic cells:
 1. **Maintenance methylation**: Addition of methyl groups occurs in a newly synthesized strand of DNA at positions opposite methylated sites on the parent strand; Maintenance activity ensures that the two daughter DNA molecules retain methylation pattern of the parent molecule
 2. **De novo methylation**: Addition of methyl groups occur at totally new positions

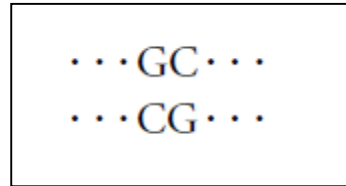
DNA methylation and gene regulation:

- **DNA methylation**- important heritable epigenetic modification of genome and involved in gene regulation
- DNA methylation is persistent throughout genomes and is missing only in regions- **CpG islands, Promoters** and possibly **Enhancers**
- This switches off eukaryotic gene expression
- In eukaryotes DNA methylation occurs at **CpG dinucleotides sequence** (p= phosphodiester bond linking Cytosine- and Guanine-containing nucleotides)
- Enzyme- **DNA methyltransferase** - mediates the transfer of a methyl group to Cytosine and generates **5-methyl-cytosine**
- Methylation of cytosine allows normal hydrogen bonding with Guanine
- High number of methylated Cytosine are usually transcriptionally inactive
- Silencing is due to direct inhibition of transcription factors binding as a result of methylated Cytosine
- Due to binding of other proteins with methyl-binding domain to the DNA

Figure 18.12

Production of 5-methylcytosine from cytosine in DNA by the action of the enzyme DNA methylase.



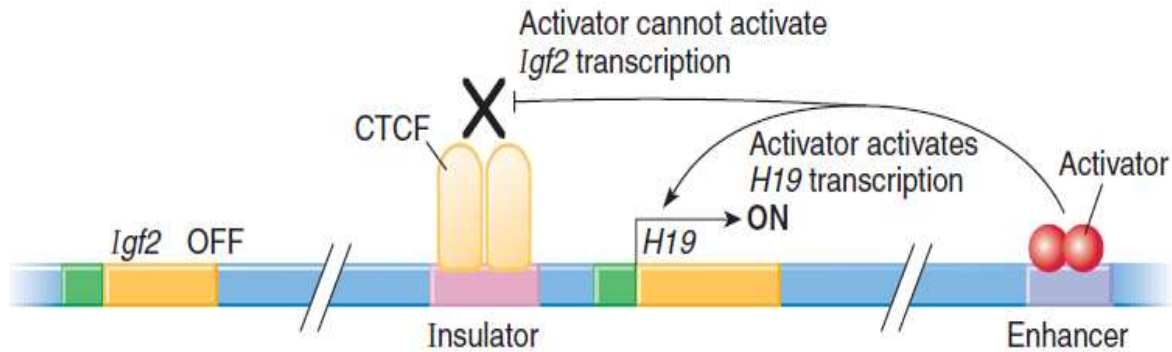
**CpG islands:**

- CpG dinucleotides are found unevenly distributed throughout the genome
- At some sites, CpG dinucleotides occur at higher frequency and referred to as **CpG islands**
- These are often associated with the transcription start sites of genes
- These islands are not methylated or have relatively low levels of methylation
- CpG island methylation of the transcription start sites is associated with stable long-term silencing (Ex. X-chromosome inactivation, imprinting, genes expressed predominantly in germ cells and some tissue-specific genes)

Regulation of transcription by methylation:

- DNA methylation is recognized by a family of proteins that contains a conserved methyl- CpG binding domain (**MBD**)
- MBDs bind to methyl group of **5mC** residues without perturbing DNA's double helical structure
- MBD-containing proteins inhibit the transcription of their bound promoter-methylated genes by recruiting protein complexes that induce the alternation of the local chromosome structure in a way that prevents the transcription of the associated gene
- Important regulatory role of DNA methylation- **Genomic Imprinting** (also called **Gametic or Parental Imprinting**)
 - Epigenetic phenomenon
 - Causes genes to be expressed in a parent-specific manner
 - In diploid organisms (like humans), the somatic cells possess two copies of each chromosome, one inherited from the father and one from the mother
 - Each autosomal gene is therefore represented by two copies (or alleles with one copy inherited from each parent at fertilization)
 - For most of the autosomal genes, expression occurs from both alleles simultaneously
 - However there are a few imprinted genes (**imprinted** meaning- gene expression occurs from only one allele) whose expression depends on whether they are inherited from the mother or from the father
 - It means the expression of an allele is dependent upon its parental origin
 - Ex. **Gene encoding insulin-like growth factor (*Igf2*) is only expressed from the father**

a) Maternal chromosome



b) Paternal chromosome

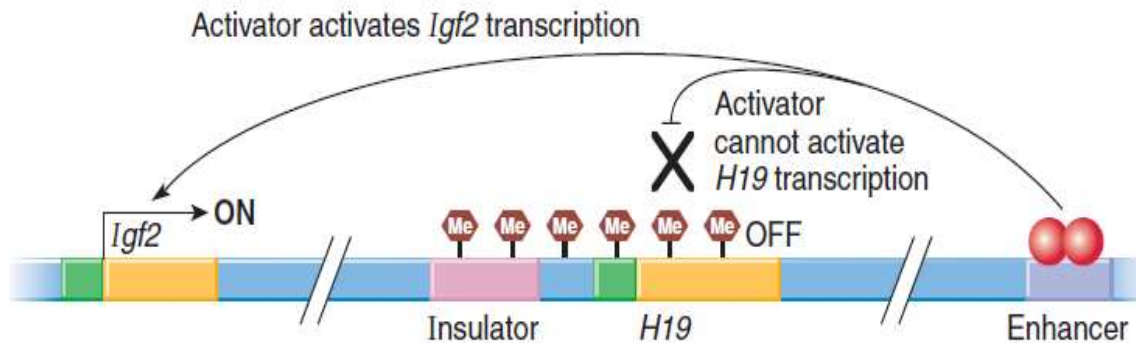


Figure 18.13

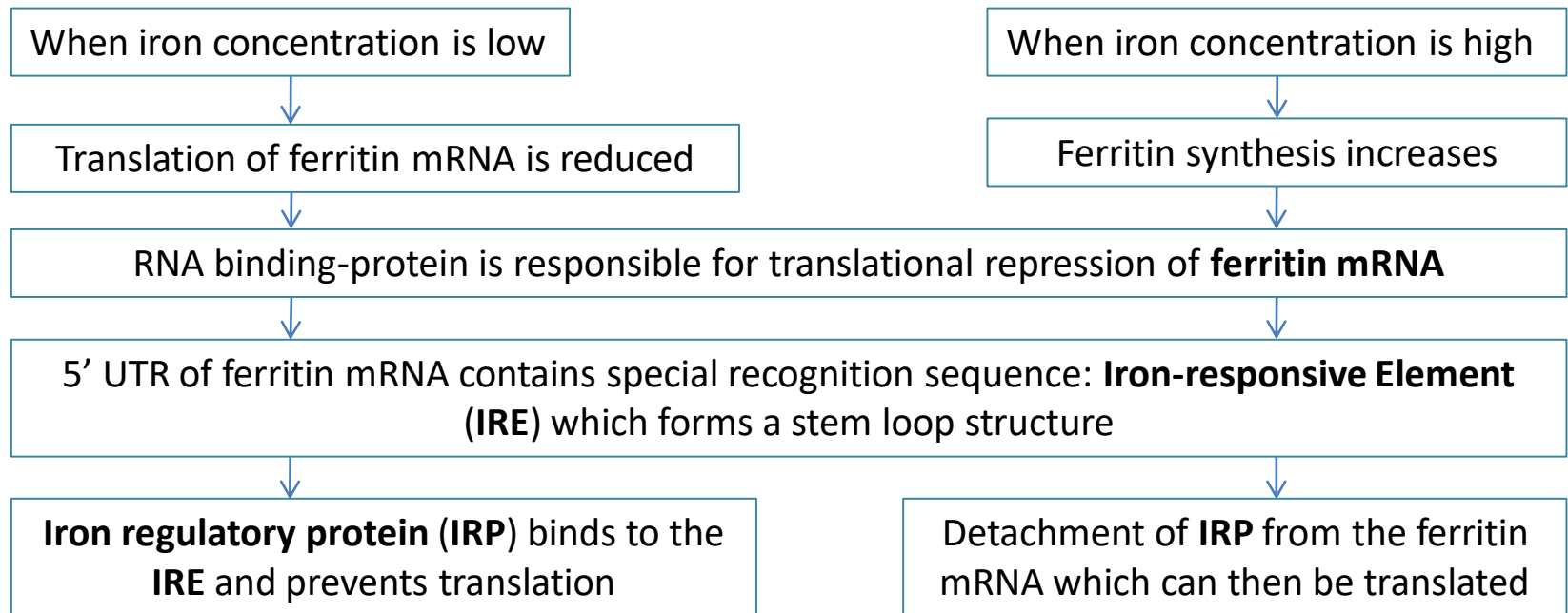
Model for the imprinting of the *Igf2* and *H19* genes.

- In some cases, only paternal allele of an imprinted gene is expressed; maternal allele is transcriptionally inactive
- For other imprinted genes, maternal allele is expressed; paternal allele is inactive
- This is a form of **Allelic Exclusion (Monoallelic Expression)**
- In such cells, the relevant gene is said to exhibit functional hemizyosity
- These epigenetic differences are conferred in the gametes, or very early in embryogenesis

- In **Germline Cells** the imprint is erased and then re-established according to the sex of the individual, i.e. in the developing sperm (during spermatogenesis), a paternal imprint is established, whereas in developing oocytes (during oogenesis), a maternal imprint is established

Post-transcriptional gene regulation:

- Regulation of gene expression occurs at the level of translation
- Less efficient and less frequent process
- Regulatory proteins may bind to DNA and either promote or repress transcription, proteins may also bind to specific sequences on a mRNA and regulate its translation
- Both positive and negative translational regulation are known
- Ex. **Translational repression- response of mRNA to iron**

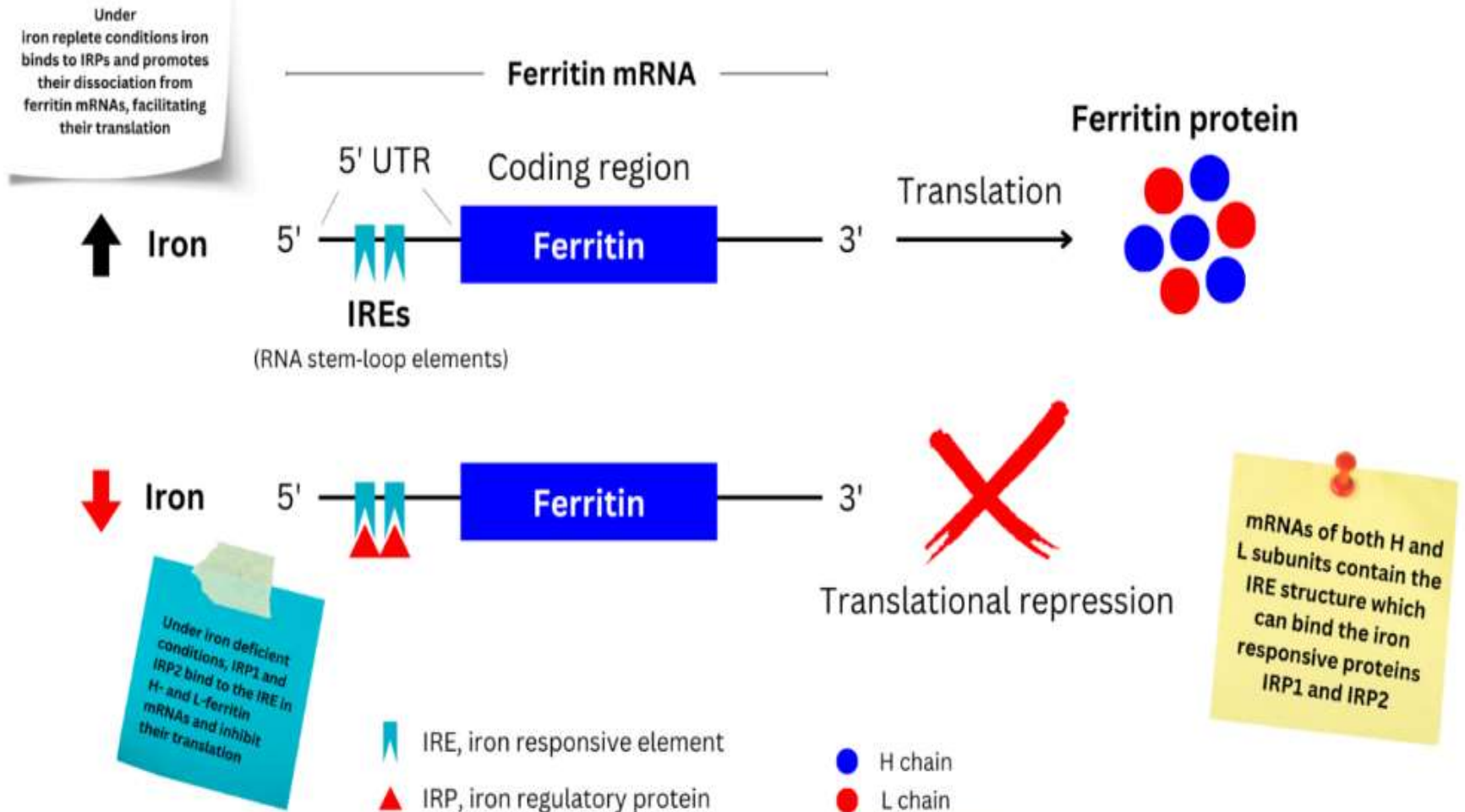


The **free iron concentration** in the cytoplasm is directly monitored by the **Iron Regulatory Proteins**. There are several IRPs, the major one is cytosolic enzyme, **Aconitase**

Protein Transferrin Receptor- transports the iron binding protein **Transferrin** is also regulated by **IRP** binding to an **IRE** present at the 3'-UTR of the mRNA, but in this case mRNA stability is affected

Post-transcriptional Control of Ferritin Expression

Iron responsive element (IRE)/iron response protein (IRP) translational repression system



RNA interference (RNAi):

- Evolutionarily conserved mechanism of gene regulation that is induced by **small silencing RNA** in a sequence-specific manner
- Fire and Mello first established this in *C. elegans*
- Also called **Post Transcriptional Gene Silencing (PTGS), Transgene Silencing, Quelling**
- Observed in all eukaryotes
- Plays important role is post-transcriptional gene regulation, transposon regulation and defending cells against viruses
- 2 types of small silencing RNA molecules are central to RNA interference:
 1. **MicroRNA (miRNA)**
 2. **Small interfering RNA (siRNA)**

miRNAs mediated RNAi:

- **miRNAs (microRNAs)**- small, non coding RNA molecules encoded in the genomes of plants, animals and their viruses
- Highly conserved, 20-25 mer RNAs appear to regulate gene expression post-transcriptionally by binding to the 3'-untranslated regions (3'-UTR) of specific mRNAs
- These small RNAs are synthesized in the nucleus as long RNA polymerase II transcripts, called **pri-miRNAs**, that are characterized by imperfect hairpin structures
- RNA polymerase III also transcribes some pri-miRNAs

siRNA mediated RNAi:

- dsRNAs are processed into siRNAs duplexes comprised of 2 (~21 nucleotides) long strands with 2 nucleotides overhangs at the 3' ends by an enzyme called **Dicer**
- Dicer (RNase III family enzyme) that functions in processing dsRNA into siRNA
- It includes an **ATPase/RNA helicase domain, catalytic RNase III domains** and **dsRNA binding domain**

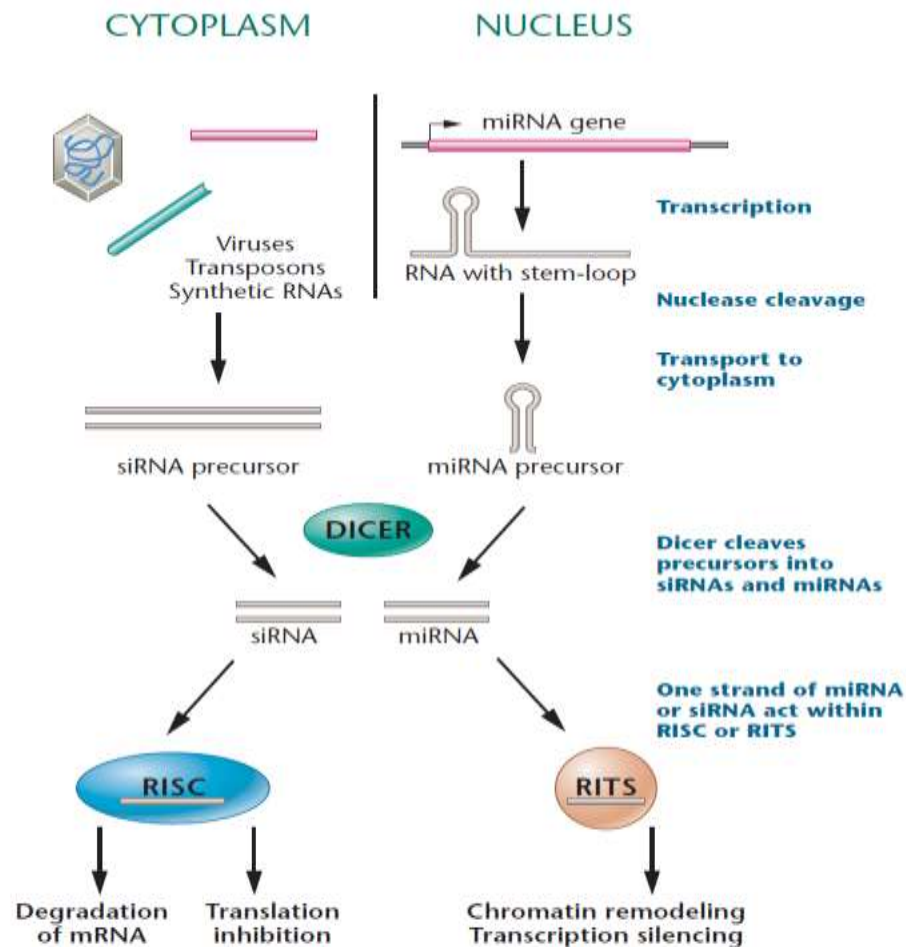


FIGURE 17-16 Mechanisms of gene regulation by RNA-induced gene silencing. The siRNA or miRNA precursors are processed into short double-stranded RNA molecules by the Dicer complex in the cytoplasm. They are then recognized by either the RISC complex or the RITS complex, and one strand is degraded. In the RNAi pathway, the RISC complex, guided by the antisense single-stranded RNA, recognizes target mRNA substrates, marking them for degradation or translation inhibition. In the transcription silencing pathway, the RITS complex acts in the nucleus, by recognizing genomic DNA that is complementary to the single strands of the miRNAs or siRNAs. The RITS complex recruits chromatin remodeling proteins that modify chromatin and repress transcription.

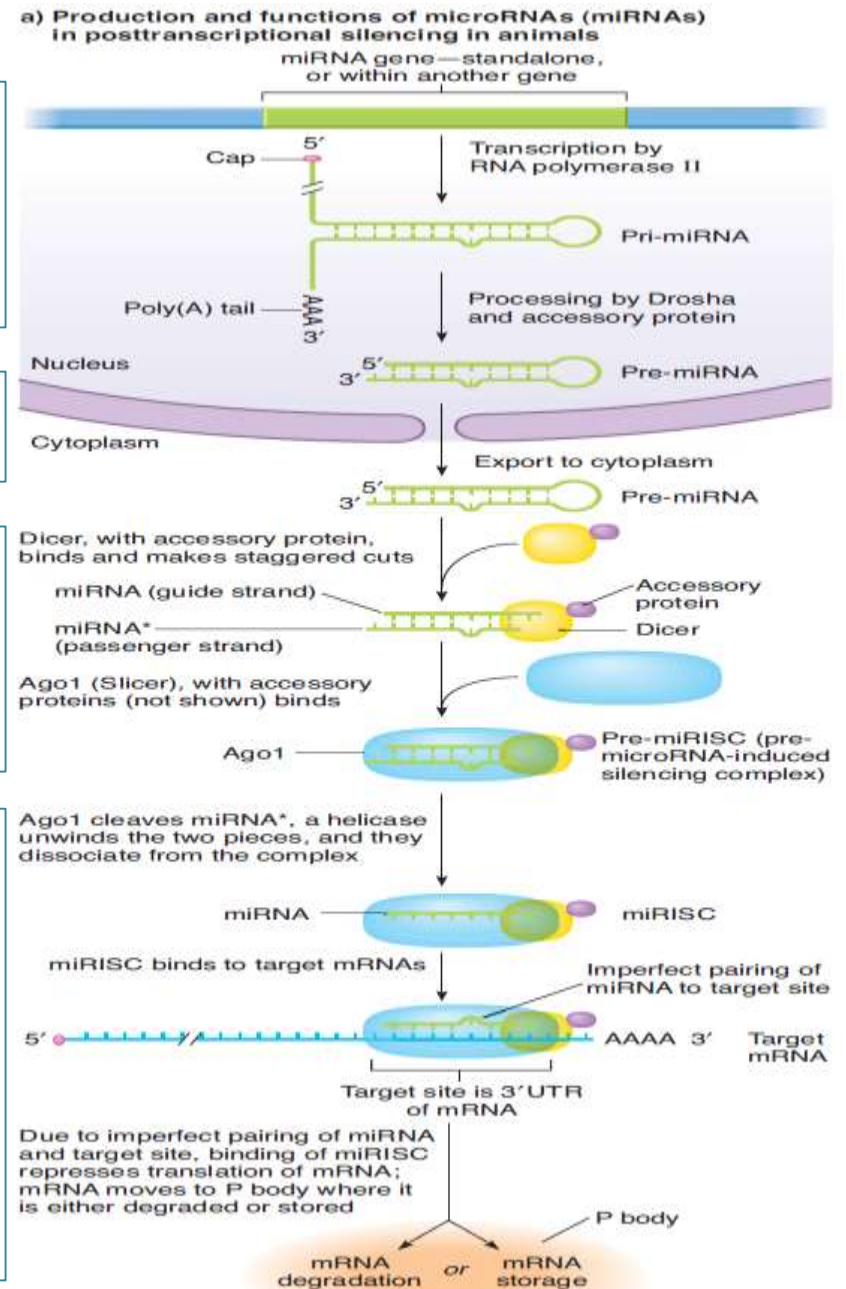
miRNAs mediated RNAi:

RNase III enzyme- Drosha acts as a dsRNA-specific endonuclease, in conjunction with a dsRNA-binding protein, called **Pasha** (in *Drosophila*) and **DGCR8** (in mammal) Processes the pri-miRNA into hairpin RNAs 70-100 nts in length, called **pre-miRNAs**

Complex of DGCR8/ pasha + enzyme Drosha= Microprocessor Complex

Pre-miRNAs also derive from introns and known as **Mirtrons** (these are not involved in processing by microprocessor complex)
Pre-miRNAs are transported to the cytoplasm. It is mediated by **Exportin-5**

In cytoplasm,
Dicer processes the pre-miRNA into miRNA duplex (**miRNA-miRNA***) and load it into the RISC
Members of the Argonaute (**Ago**) protein family are central to RISC function
These are needed for miRNA-induced silencing and contain 2 conserved RNA binding domains:
PAZ domain- bind the single stranded 3' end of the mature miRNA
PIWI domain- structurally resembles Ribonuclease-H



miRNA biogenesis:

In plants,

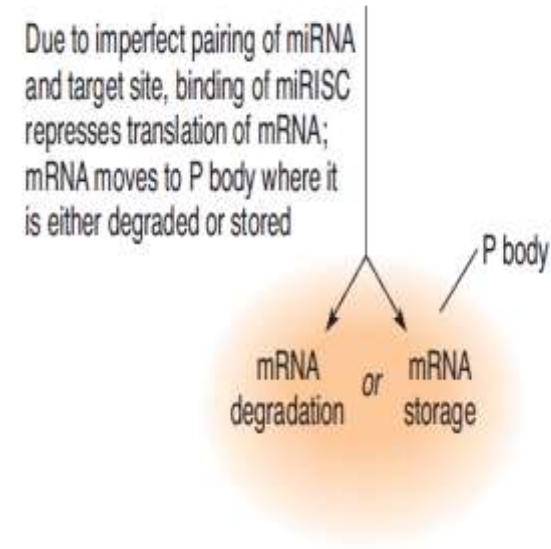
Single enzyme **DCL1** present inside nucleus fills the roles of both Drosha and Dicer, converts pri-miRNA to mature miRNA

Before **miRNA-miRNA*** duplexes are transported out of the nucleus, its 3' overhangs are methylated by a **S- adenosyl methionine- dependent methyltransferase** called **HEN1**

In animal **miRNAs** are able to recognize their target mRNAs by seed sequence present at the 5' end

The seed sequence of a miRNA is defined as the first 2-8 nucleotides starting from the 5' end and counting toward the 3' end

- Gene silencing may occur either via **mRNA degradation** or **preventing mRNA from being translated**
- miRNA function via base-pairing with complementary sequences within target mRNA molecules
- If there is complete complementation between the miRNA and target mRNA sequence, **Ago** can cleave the mRNA and lead to direct **mRNA degradation**
- If there is not complete complementation, silencing is achieved by preventing translation
- Plant miRNA usually have perfect/ near perfect pairing with their target mRNA and induce gene repression through degradation of their target mRNAs



siRNA mediated RNAi:

Dicer+ dsRNA binding protein= RISC loading complex

Loading of **RNA duplex** into **RISC**

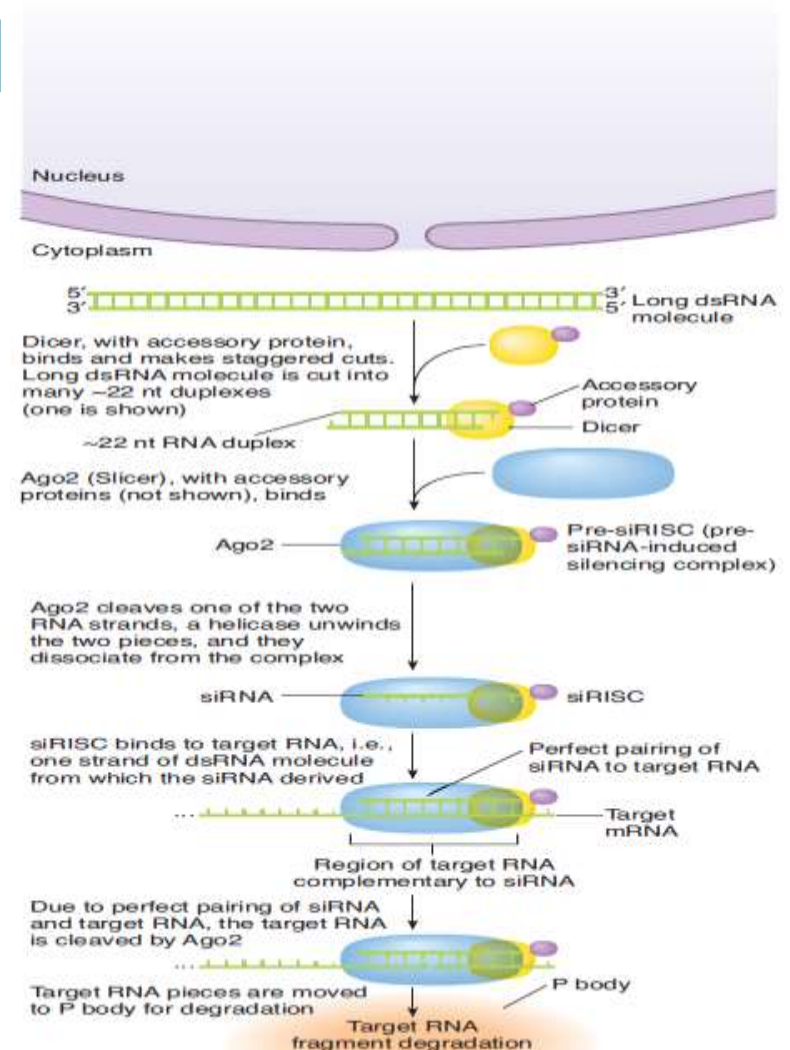
siRNA provide target specificity to RISC through base pairing of the guide strand with the target mRNA

siRNA are fully complementary to their targets

Only one of the two strands (guide strand) directs the gene silencing
The other anti-guide strand/ passenger strand is degraded during RISC activation

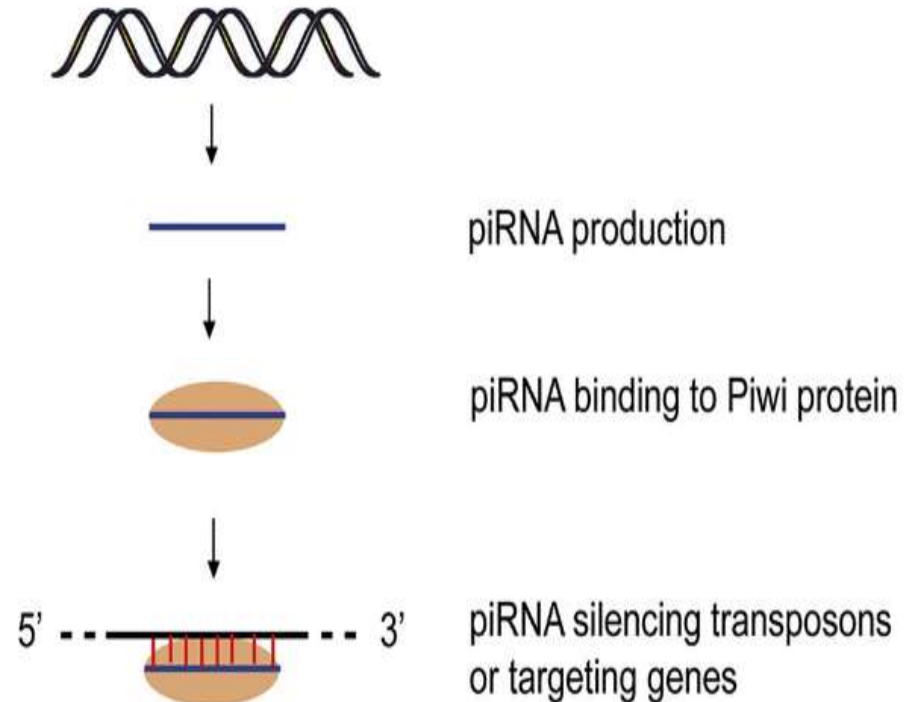
The active components of an RNA-induced silencing complex (RISCs) are endonucleases called **Argonaute proteins** which cleave the target mRNA strand complementary to their bound siRNA

b) Production and functions of short interfering RNAs (siRNAs) in posttranscriptional silencing in animals



piRNA:

- PIWI-interacting RNAs
- These are found in metazoans
- Most recently discovered class of longer small RNAs
- Binds to the Piwi clade of Argonaute proteins
- Derived from ssRNA precursor and are made without a dicing step (dicer independent)
- They are mostly antisense, but a small fraction is in the sense orientation
- They are 2'-O-methylated at their 3' termini
- They direct cleavage of transposon mRNA or chromatin modification at transposon loci



<https://www.cell.com/molecular-therapy-family/nucleic-acids/fulltext/S2162-2531%2820%2930149-9>

siRNA	miRNA
siRNAs are processed from dsRNA precursors made up of two distinct strands of perfectly base-paired RNA	miRNA originated from a single, long transcript that forms imperfectly base-paired hairpin structures
These are originally identified as intermediates in the RNAi pathway after induction by exogenous dsRNA (endogenous source of siRNA have also been recognized)	These are discovered through their critical roles in development and cellular regulation Has been recognized as being of endogenous origin
siRNAs derived from long dsRNA precursors	miRNAs processed from stem-loop precursors with incomplete double-stranded character
They typically base-pair perfectly and induce mRNA cleavage only in a single, specific target	They have an incomplete base pairing to a target and inhibit the translation of many different mRNAs with similar sequences

Types		Organisms	Length	Features
miRNA		Animals, plants, protists	20-25	Dicer/ Drosha dependent
siRNA	Exo-siRNA	Animals, plants, fungi, protists	~21	Dicer dependent
	Endo-siRNA	Animals, plants, fungi, protists	~21	Dicer dependent
piRNA		Metazoans	24-30	Dicer independent

References:

1. Benjamin A Pierce, Genetics, A conceptual approach, Page no. 436-465
2. Snustad, Simmons (2012), Principles of genetics, John Wiley& Sons, Inc, Sixth edition, Page no. 504-526
3. Peter J Russell, iGenetics A molecular approach, 3rd edition, Page no.491-512
4. Klug cummins, Spencer, Palladino (2012), Concepts of Genetics, Pearson Publication, 10th Edition, Page no.403-445